Shh and ZRS enhancer colocalisation is specific to the zone of polarizing activity

Development Advance Online Articles. First posted online on 11 July 2016 as 10.1242/dev.139188
Access the most recent version at http://dev.biologists.org/lookup/doi/10.1242/dev.139188

Shh and ZRS enhancer co-localisation is specific to the zone of polarizing activity

Iain Williamson, Laura A. Lettice, Robert E. Hill*, Wendy A. Bickmore*

MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, Crewe Road, Edinburgh EH4 2XU, UK

* Correspondence to: Wendy.Bickmore@igmm.ed.ac.uk or Bob.Hill@igmm.ed.ac.uk

6 key words: 5C, chromosome loop, enhancer, limb development, super-resolution microscopy,

Abstract

Limb-specific *Shh* expression is regulated by the (~1 Mb distant) ZRS enhancer. In the mouse, limb bud restricted spatiotemporal *Shh* expression occurs from ~E10-E11.5 at the distal posterior margin and is essential for correct autopod formation. Here, we have analysed the higher-order chromatin conformation of *Shh* in expressing and non-expressing tissues, both by fluorescence in situ hybridisation (FISH) and by chromosome conformation capture (5C). Conventional and super-resolution light microscopy identified significantly elevated frequencies of *Shh*/ZRS co-localisation only in the *Shh* expressing regions of the limb bud, in a conformation consistent with enhancer-promoter loop formation. However, in all tissues and developmental stages analysed, *Shh*-ZRS spatial distances were still consistently shorter than those to a neural enhancer located between Shh and ZRS in the genome. 5C identified a topologically associating domain (TAD) over the *Shh*/ZRS genomic region and enriched interactions between *Shh* and ZRS throughout E11.5 embryos. *Shh*/ZRS co-localisation, therefore, correlates with the spatiotemporal domain of limb bud-specific *Shh* expression, but close *Shh*/ZRS proximity in the nucleus occurs regardless of whether the gene or enhancer is active. We suggest that this constrained chromatin configuration optimises the opportunity for the active enhancer to locate and instigate *Shh* expression.

Introduction

Chromatin-looping is a popular model by which very long-range enhancers can communicate with their target gene promoter (Benabdallah and Bickmore, 2015), however the relationship of loop formation and gene activation remains unclear. It has been suggested that enhancer-target gene contacts are preformed and present in tissues even where the target gene is not activated (Montavon et al., 2011; Ghavi-Helm et al., 2014) . However, other reports indicate enhancer-gene looping is spatially and temporally restricted to cells where the target gene is active. This includes in the developing mouse limb, where elevated levels of co-localisation of the global control region (GCR) and its target 5′HoxD genes is only seen in the cells of the distal posterior portion of the E10.5 limb bud (Williamson et al., 2012).

The complex spatiotemporal gene regulatory circuit in the developing limb is a rich system in which to study the activity of distal regulatory elements and their mechanisms of action. The sonic hedgehog gene (*Shh*), encodes a morphogen that directs cell fate during organogenesis. Limb-specific expression of *Shh* is regulated by the ZRS enhancer positioned within an intron of *Lmbr1* ~1 Mb away at the opposite end of a large gene desert (Lettice et al., 2002; Lettice et al., 2003) (Figure 1A). The ZRS has a functional role in directing spatiotemporal *Shh* expression restricted to a region of the distal posterior mesenchyme of the limb bud known as the zone of polarizing activity (ZPA) (Saunders and Gasseling, 1968; Riddle et al., 1993). Limb-specific *Shh* expression is abrogated upon deletion of ZRS (Sagai et al., 2005), whereas point mutations across the 780-bp conserved sequence of the enhancer can induce anterior, ectopic *Shh* expression and can cause preaxial polydactyly (Lettice et al., 2003; Sagai et al., 2004; Lettice et al., 2008), triphalangeal thumb (Furniss et al., 2008) or Werner mesomelic syndrome (VanderMeer et al., 2014). Duplications, and even triplication, of the ZRS have been associated with severe forms of polysyndactyly: triphalangeal thumb-polysyndactyly syndrome and Haas type (syndactyly type IV) polysyndactyly (Klopocki et al., 2008; Sun et al., 2008; Wieczoreck et al., 2010).

Previously, fluorescence in situ hybridisation (FISH) and chromosome conformation capture (3C) (Amano et al., 2009) have reported increased associations between *Shh* and ZRS in E10.5 limb buds compared with other tissues. However, no significant difference in gene/enhancer co-localisation was detected between the ZPA and distal anterior tissue – where *Shh* is not normally expressed, or indeed in ZPA cells between wild-type and embryos with a deletion of the ZRS. This would be consistent with a model of pre-formed enhancer-

gene contacts. In contrast, FISH has revealed a significant decrease in *Shh*/ZRS co-localisation in E11.5 ZPA tissue from mouse embryos with a ZRS mutation which decreases ZRS long-range activity (Lettice et al., 2014), suggesting that ZRS/*Shh* juxtaposition is directly linked to *Shh* activation.

We have previously combined FISH and 3C carbon copy (5C) to elucidate the role of chromatin conformation in the long-range regulation of the 5′ *Hoxd* genes during distal limb bud development (Williamson et al., 2012; Williamson et al., 2014). Here we combined these methods to characterise the *Shh* locus in tissue sections – including those derived from three discrete developmental stages of mouse limb bud development. Spatial proximity of *Shh* and ZRS, as inferred indirectly from enriched 5C interactions, was identified throughout E11.5 embryos, and 5C data confirmed that *Shh* and its known enhancers form a compact regulatory chromatin domain. However, using super-resolution microscopy we show that, despite *Shh* and ZRS being proximal to one another in the nucleus in all tissue types and temporal stages analysed, high levels of *Shh*/ZRS co-localisation occurs only in ZPA cells at the time of *Shh* activation. Comparison between Shh/ZRS distances and those between either *Shh* or ZRS and an intervening genomic locus are consistent with the formation of a chromatin loop between the active gene and enhancer.

Materials and Methods

FISH

For 3D FISH, E10.5, E11.5 and E14.5 embryos from CD1 mice were collected, fixed, embedded, sectioned and processed as previously described (Morey et al., 2007), except that sections were cut at 6 μm. Fosmid clones (Figure 1A, Table S1) were prepared and labelled as previously described (Morey et al. 2007). Between 160-240 ng of biotin- and digoxigenin-labeled fosmid probes were used per slide, with 16-24 μg of mouse Cot1 DNA (Invitrogen) and 10 μg salmon sperm DNA.

Image analysis

For 3D analysis of tissue sections by conventional microscopy, slides were imaged with a Hamamatsu Orca AG CCD camera (Hamamatsu Photonics (UK) Ltd, Welwyn Garden City, UK), Zeiss Axioplan II fluorescence microscope with Plan-neofluar or Plan apochromat objectives, a Lumen 200W metal halide light source (Prior Scientific Instruments, Cambridge, UK) and Chroma #89014ET single excitation and emission filters (Chroma Technology Corp., Rockingham, VT) with the excitation and emission filters installed in Prior motorised filter wheels. A piezoelectrically driven objective mount (PIFOC model P-721, Physik Instrumente GmbH & Co, Karlsruhe) was used to control movement in the z dimension. Hardware control, image capture and analysis were performed using Volocity (Perkinelmer Inc, Waltham, MA). Images were deconvolved using a calculated point spread function with the constrained iterative algorithm of Volocity (Perkinelmer Inc, Waltham, MA). Image analysis was carried out using the Quantitation module of Volocity (Perkinelmer Inc, Waltham, MA).

SIM imaging

Images were acquired using Structured Illumination Microscopy (SIM) performed on an Eclipse Ti inverted microscope equipped with a Nikon Plan Apo TIRF objective (NA 1.49, oil immersion) and an Andor DU-897X-5254 camera. Laser lines 405, 488 and 561 nm were used. Step size for *z* stacks was set to 0.120 μm, which is well within the Nyquist criterion. For each focal plane, 15 images (5 phases, 3 angles) were captured with the NIS-Elements software. SIM image processing and reconstruction were carried out using the N-SIM module of the NIS-Element Advanced Research software. Image analysis was carried out using the Quantitation module of Volocity (Perkinelmer Inc, Waltham, MA) with *x* and *y* binning resolution of 32 nm.

3C library preparation

Limbs from ~70 E11.5 embryos, 3 E11.5 embryos with the limbs and heads removed, and the heads of 3 E11.5 embryos were collected in 15 ml tubes with enough PBS to cover them and to dissociate the cells by repeated pipetting with enlarged tip ends. Cells were fixed with 1% formaldehyde for 10 min at room temperature (r.t.). Crosslinking was stopped with 125 mM glycine, for 5 min at r.t. followed by 15 min on ice. Cells were centrifuged at 400 *g* for 10 min at 4°C, supernatants removed and cell pellets flash frozen on dry ice.

Cell pellets were treated as previously described (Dostie and Dekker 2007; Ferraiuolo et al., 2010; Williamson et al., 2014). HindIII-HF (NEB) was the restriction enzyme used to digest the crosslinked DNA.

5C primer and library design

5C primers covering the *Usp22* (mm9, chr11: 60,917,307-61,003,268) and *Shh* regions (mm9, chr5: 28,317,087-30,005,000) were designed using 'my5C.primer' (Lajoie et al. 2009) and the following parameters: optimal primer length of 30 nt, optimal TM of 65°C, default primer quality parameters (mer:800, U-blast:3, S-blasr:50). Primers were not designed for large (>20 kb) and small (<100 bp) restriction fragments, for low complexity and repetitive sequences, or where there were sequence matches to >1 genomic target. The *Usp22* region was used to assess the success of each 5C experiment but was not used for further data normalization or quantification.

The universal A-key (CCATCTCATCCCTGCGTGTCTCCGACTCAG-(5C-specific)) and the P1-key tails ((5C-specific)-ATCACCGACTGCCCATAGAGAGG) were added to the Forward and Reverse 5C primers, respectively. Reverse 5C primers were phosphorylated at their 5′ ends. An alternating design consisting of 365 primers in the *Shh* region (182 Forward and 183 Reverse primers) was used. Primer sequences are listed in Table S6.

5C library preparation

5C libraries were prepared and amplified with the A-key and P1-key primers as previously described (Fraser et al. 2012). Briefly, 3C libraries were first titrated by PCR for quality control (single band, absence of primer dimers, etc.), and to verify that contacts were amplified at frequencies similar to those usually obtained from comparable libraries (same DNA amount from the same species and karyotype) (Dostie and Dekker 2007,, Dostie, et al. 2007, Fraser, et al. 2010). We used 1 - 10 μg of 3C library per 5C ligation reaction.

5C primer stocks (20 μM) were diluted individually in water on ice, and mixed to a final concentration of 2 nM. Mixed diluted primers (1.7 μl) were combined with 1 μl of annealing buffer (10X NEBuffer 4, New England Biolabs Inc.) on ice in reaction tubes. 1.5 μg salmon testis DNA was added to each tube, followed by the 3C libraries and water to a final volume of 10 μl. Samples were denatured at 95°C for 5 min, and annealed at 55°C for 16 hours. Ligation with Taq DNA ligase (10 U) was performed at 55°C for one hour. One tenth (3 μl) of each ligation was then PCR-amplified individually with primers against the A-key and P1-key primer tails. We used 26 cycles based on dilution series showing linear PCR amplification within that cycle range. The products from 3 to 5 PCR reactions were pooled before purifying the DNA on MinElute columns (Qiagen).

5C libraries were quantified by bioanalyser (Agilent) and diluted to 26 pmol (for Ion PGM™ Sequencing 200 Kit v2.0). One microlitre of diluted 5C library was used for sequencing with an Ion PGM™ Sequencer. Samples were sequenced onto Ion 316™ Chips following the Ion PGM™ Sequencing 200 Kit v2.0 protocols as recommended by the manufacturer (Life Technologies™).

5C data analysis

Analysis of the 5C sequencing data was performed as previously described (Berlivet et al., 2013). Sequencing data was processed through a Torrent 5C data transformation pipeline on Galaxy (https://main.g2.bx.psu.edu/). Data was normalized by dividing the number of reads of each 5C contact by the total number of reads from the corresponding sequence run. All scales shown correspond to this ratio multiplied by 10^3. For each experiment the number of

total reads, and of used reads, is provided in Table S7. The unprocessed heatmaps of the normalized 5C datasets can be found in Figure S4. 5C datasets are uploaded to the Gene Expression Omnibus (GEO) website (http://www.ncbi.nlm.nih.gov/geo/) under Accession number: GSE79947.

Results

Increased co-localisation of ZRS with *Shh* in the limb ZPA at E10.5 and E11.5

Previous analyses of the chromatin dynamics involved in the long-range regulation of *Shh* by ZRS has produced contradictory results, which could be due to the different temporal stages of development assayed (Amano et al., 2009; Lettice et al., 2014). To resolve this issue we carried out FISH on whole mouse embryo sections that include posterior and anterior forelimb tissue from E10.5, E11.5 and E14.5 developmental stages (Figure 1B). *Shh* is expressed within the ZPA at the two earlier stages but is switched off in the limb by E14.5 (Riddle et al., 1993). We compared inter-probe distances (Figure S1A shows representative images) and co-localisation frequencies (Figure 1C; left) between the gene and enhancer in mesenchymal tissue across the anterior-posterior axis of the distal forelimb bud. In addition proximal limb tissue and the adjacent flank where *Shh* is not expressed were compared.

By conventional wide-field deconvolution microscopy, the proportion of co-localised (<200 nm) *Shh* and ZRS probe pairs in ZPA cells at stages when *Shh* is expressed (E10.5, E11.5), was significantly higher (35%) than in inactive limb regions and the flank (distal anterior $p < 0.05$, proximal $p < 0.01$, flank $p < 0.001$ (Figure 1C, Table S2)). By E14.5 there is no *Shh* expression in the limb (expression ends between E11.5-E12.5), and the *Shh*-ZRS co-localisation frequency in distal posterior cells is significantly reduced, compared to E10.5 and E11.5 ZPA (E10.5 $p = 0.006$, E11.5 $p = 0.01$) (Figure 1D). At this later stage, differences in co-localisation frequencies between the distal posterior forelimb region (~20%) and the other limb regions and the flank mesoderm are also no longer detected (Figure 1C).

Dpp6 is located the same linear genomic distance away from *Shh* as the ZRS, but in the other direction and outside of the *Shh* regulatory domain (Figure 1A). In contrast to the spatial proximity of *Shh*-ZRS, *Shh* and *Dpp6* are predominantly located > 400 nm apart (co-

localisation frequency < 5%) (Figure 1C; right) for all tissues and developmental stages examined.

The greater co-localisation of the active enhancer (ZRS) with its target gene (*Shh*) in the ZPA at E10.5 is similar to what we reported for these loci at E11.5 (Lettice et al. 2014) and to the co-localisation frequency of *Hoxd13* and its GCR enhancer in distal posterior expressing limb tissue and cell lines at E10.5 (Williamson et al., 2012; Williamson et al., 2014). These differences in three-dimensional chromatin conformation between active and inactive tissues contradicts the previous report suggesting an equivalent rate of *Shh*-ZRS co-localisation on both sides of the distal limb field at this developmental stage (Amano et al., 2009).

Super-resolution imaging identifies *Shh*/ZRS co-localisation of most alleles in the ZPA

The data in Figure 1 are consistent with active gene-enhancer co-localisation during long-range regulation. From the images of tissue sections from the three developmental stages, acquired by conventional light-microscopy, it was apparent that *Shh* and ZRS are consistently very close in the nucleus, with differences in spatial distance frequently down to the signal centroids being in different layers of the *z* stack – the dimension with the lowest spatial resolution in the microscope. We therefore re-analysed the tissue sections containing E10.5 and E11.5 distal anterior and posterior (ZPA) cells by structured illumination microscopy (3D-SIM) (Figure 2A). This technique doubles the resolution limit in all dimensions (Toomre & Bewersdorf, 2010) and has previously been combined with 3D-FISH (Nora et al., 2012; Patel et al., 2013).

The greater resolution afforded by 3D-SIM, particularly for the *z* (depth) dimension (120 nm compared to 200 nm in conventional widefield microscopy), not only confirmed the difference in *Shh*/ZRS co-localisation frequency between ZPA and distal anterior limb bud but also suggests that conventional microscopy does not fully capture the proportion of co-localised *Shh*/ZRS probe pairs, especially in the *Shh*-expressing tissues where it now peaks at 79% (Figure 2B, Table S3). These data suggest that a substantial proportion of *Shh*/ZRS probe pairs with signal centroids not in the same plane of the *z* stack, that have been categorised as adjacent (between 200 nm and 400 nm apart (Figure S1)) due to the low *z* dimension resolution afforded by conventional widefield microscopy, are indeed co-localised

in ZPA cells. At both temporal stages the anterior/posterior differences in *Shh*/ZRS co-localisation frequency were highly significant (E10.5 $p = 0.0002$, E11.5 $p = 0.0001$). Due to variation in fluorescent probe signal strength between alleles in the same nucleus, and the limited number of *z* stack planes imaged per nucleus by SIM to minimize fluorochrome bleaching, generally less than half of all probe pairs measured from each tissue in Figure 2 are from both alleles of the same nucleus.. However, for cells where both alleles could be measured, ZRS/*Shh* colocalisation at both occurred in 33% (E10.5) and 59% (E11.5) of ZPA cells. The proportion of ZPA cells with only one co-localised allele was 56% (E10.5) and 31% (E11.5). Only around a tenth of ZPA cells at both temporal stages had no co-localising alleles, compared to a third of distal anterior cells. By comparing conventional and SIM data for the *Shh*/ZRS probe pair in anterior and posterior tissues at two developmental stages, we show that median inter-probe distances in distal anterior limb tissues are very similar when measured by either technique (conventional = 250 nm, SIM: E10.5 = 246 nm, E11.5 = 275 nm); whereas, these are significantly different for ZPA cells (E10.5: conventional = 221 nm, SIM = 160 nm, $p = 0.01$; E11.5: conventional = 241 nm, SIM = 136 nm, $p < 0.0001$) (Figure 2C, Tables S4 & S5).

The *Shh*/ZRS regulatory domain is compact in expressing and non-expressing tissue

Long-range gene/enhancer co-localisation is often depicted as a looping out of the intervening chromatin fibre (Williamson et al., 2011). Our previous work on the HoxD locus implicated a gross compaction of the regulatory region, rather than a simple loop with extrusion of the intervening chromatin, upon activation of *Hoxd13* by the long-range (~250-kb) limb-specific GCR enhancer (Williamson et al., 2012; Williamson et al., 2014). We therefore used 3D FISH and conventional wide-field deconvolution microscopy to measure the spatial distances between either *Shh* or the ZRS, and the SBE4 enhancer that drives *Shh* expression in the forebrain (Figure 3A) (Jeong et al., 2006). SBE4 is located midway through the gene desert separating Shh and ZRS (Figures 1A). If the entire genomic region between the gene and the limb enhancer forms a loop then *Shh*-SBE4 and SBE4-ZRS distances should be greater than those between *Shh* and ZRS.

At both temporal stages (E10.5 and E11.5) when *Shh* is active in the distal posterior limb mesenchyme, but not at E14.5, *Shh* is closer to SBE4, and Shh/SBE4 co-localisation

frequencies are higher, compared to the other tissues analyzed (Figure 3B & C, S2A & B, and Table S5). These data suggest that the genomic region between *Shh* and the ZRS is folded into a compact chromatin domain, which is at its most compact in distal posterior *Shh*-expressing cells. However, what is also apparent is that the spatial distances between *Shh* and the ZRS are less than those between either *Shh*-SBE4 or SBE4-ZRS in most expressing and non-expressing tissues (Figure S2C). These differences are significant for most tissues analysed and intriguingly is particularly apparent at E14.5, well past the stage of limb-specific *Shh* activity and therefore could be indicative of a constitutive chromatin conformation.

Topography of the *Shh* regulatory domain is maintained throughout the E11.5 embryo

Using FISH we could only infer the conformation of the Shh regulatory domain from the spatial relationships of three genomic loci across the *Shh*-ZRS region. To gain a more complete view of the locus, we used 5C to determine the frequency of cross-linked interactions captured between sequences in the ~1.7 Mb region from *Irsig1* ~400kb 3′ of *Shh* to *Ube3c* ~350kb beyond the ZRS (Figure 1A) in dissected whole fore- and hindlimb buds from ~70 E11.5 embryos (x2 replicates) (Figures 4A, left-hand heatmap, S3A and S4). We were unable to dissect cells suitable for 5C specifically from the ZPA. Three interaction domains can be identified; with the middle topologically associated domain (TAD) containing *Shh* and its entire known regulatory elements with the boundaries located 3′ of *Rbm33* and within the 5′ end of *Lmbr1*. This *Shh* regulatory TAD corresponds well with that identified by Hi-C in mouse ESCs (Dixon et al., 2012)

In limb cells 5C cross-linked interactions are enriched between genomic fragments across the *Shh* and ZRS loci (Figures 4A, left-hand heatmap, S3A and S4). The general spatial proximity of *Shh* and the ZRS detected by FISH and inferred from enriched 5C interaction frequencies in expressing and non-expressing tissues suggests that this conformation is constitutive. To determine whether the high cross-linking efficiency of *Shh* and ZRS identified in E11.5 limb buds can also be detected in tissues where the ZRS is not active we carried out 5C on cells derived from the bodies and heads of E11.5 embryos. Even with the vast majority of these cells not expressing *Shh*, high read frequencies between *Shh* and ZRS were captured (Figures 4A, middle and right-hand heatmaps, and S4) and the same TAD structures could be discerned as seen in limb tissue.

To examine more closely the regions probed by FISH (*Shh*, SBE4 and ZRS) we generated "virtual 4C" plots from the 5C data (Figures 4B, S3B) (Williamson et al., 2014). From the viewpoint of *Shh*, overall interaction frequencies with the rest of its regulatory domain is similar in limb-, body- and head-derived tissues, and are not substantially higher than those extending into the adjacent TAD 3′ of *Shh* (Figures 4B compare the top track with the track that profiles SBE4 located in the middle of a TAD). Highest interaction frequencies for *Shh*, apart from genomic regions immediately adjacent, are with regions within the neighbourhood of ZRS (limb-specific high interactions with a loci within the gene desert that does not contain any known regulatory elements was not identified in the limb replicate data (Figure S3B)). ZRS has reciprocal enriched interactions with the *Shh* region (Figures 4B, bottom track). However, these are not detectably higher in limb than in the embryonic body or head.

Discussion

Activation of *Shh* in the limb bud is accompanied by co-localisation with the ZRS

Using 3D-FISH and super-resolution imaging, we provide compelling evidence that co-localisation (<200 nm) between *Shh* and the ZRS enhancer is associated with *Shh* expression in the ZPA region of the distal posterior forelimb bud, to an extent not seen in control tissues, including the limb bud after *Shh* expression has ceased at E14.5 (Figures 1 and 2). The co-localisation frequencies detected by super-resolution microscopy rise to almost 80% at E11.5 – suggesting that the vast majority of *Shh* alleles in the ZPA are juxtaposed to the ZRS located 1Mb of genomic distance away. Analysis of the FISH images by either conventional wide-field, or structured illumination microscopy, showed a significantly higher gene-enhancer co-localisation frequency in the ZPA than in nuclei from the distal anterior region of the same limb buds (Figures 1 and 2). This anterior-posterior difference in chromain folding is consistent with our previous analysis for *Shh*-ZRS in E11.5 fore- and hindlimbs (Lettice et al. 2014) and is similar to the preferential co-localisation of *Hoxd13*-GCR in E10.5 distal posterior limb buds (Williamson et al., 2012). Like *Shh*, *Hoxd13* expression is restricted to the posterior margin of the distal limb bud at this stage. These data, however, contradict previously published work that could not identify signficant differences in *Shh*-

ZRS proximity between the *Shh*-expressing ZPA and distal anterior cells (Amano et al., 2009). Those data were derived from single cell suspensions of dissected tissue from specific points across the distal limb bud whereas our data are from sections cut through whole embryos; therefore cell/tissue preparation may be a factor in discrepencies between the data sets.

Shh and its regulatory elements are located within a compact chromatin domain

Long-range interactions between genes and *cis*-regulatory elements are usually described as loops, which should be visualized as a coming together of the two loci to the exclusion of the intervening chromatin (Williamson et al. 2011; Fraser et al. 2015). Indeed a looping mechanism in distal limb could be inferred from the shorter inter-probe distances between *Shh* and ZRS, than for either of these probes with the forebrain SBE4 enhancer – even though the latter is located midway between *Shh* and ZRS on the linear chromosome (Figure S2C). To our knowledge this apparent *Shh*/ZRS chromatin 'loop' is the first to be identified by FISH.

However, the *Shh*-ZRS distances are shorter than distances to SBE4 not only in the ZPA but also in anterior limb and in E14.5 tissues when the ZRS is no longer active. But, *Shh*-ZRS co-localisation frequencies are not significant in those tissues. Another interpretation of these data is that the *Shh* regulatory domain (Figures 4A, S3A, S4) is maintained in a tightly folded chromatin conformation where *Shh* and the ZRS are generally proximal in nuclear space. That the *Shh*-containing TAD is indeed compact can be discerned from the frequency distribution graphs which show that most *Shh*/ZRS, *Shh*/SBE4 and SBE4/ZRS probe pairs are adjacent (200 – 400 nm) or co-localised (<200 nm), with median interprobe distances of between 220 – 345 nm for most tissues and developmental stages analysed (Figures 1B, S2A & B; Table S5). This is consistent with our 5C analysis of E11.5 limb bud, body and head cells which suggests that the *Shh* regulatory region forms a constitutive self-interacting domain – the *Shh* TAD has also been identified in ES cells by Hi-C (Dixon et al. 2012). These data show somewhat enriched interactions between cross-linked DNA fragments from the genomic regions containing *Shh* and ZRS (Figures 4, S3, and S4), but in all analysed tissues/cell types. The very high co-localisation frequencies that we see by microscopy between ZRS and *Shh* in the distal posterior limb at stages of *Shh* expression are not reflected in elevated interactions captured by 5C. Similarly, the increased compaction of

the intervening genomic region in ZPA cells inferred from FISH analysis of distances to the neural SBE4 enhancer could not be identified by 5C. We do not know whether this is because the *Shh* expressing (ZPA) cells do not present at a high enough proportion of cells in the dissected limb buds, or because the spatial proximities of *Shh* and ZRS, and the ZPA-specific chromatin domain is not well captured by chromosome conformation methods (Belmont, 2014). Conversely, our previous analysis comparing 5C and FISH has highlighted that spatial proximity should not always be inferred from enriched cross-linked interactions between 3C fragments (Williamson et al., 2014).

Facilitating gene regulation by enhancer – promoter proximity

Here, we have shown that local chromatin conformation maintains spatial proximity of *Shh* with the regulatory domain containing its enhancers – including the limb-specific enhancer ZRS – in a variety of cell types – not just those expressing *Shh*. If the physical interaction of active enhancers and their target gene promoters is essentially a stochastic process, their constitutive relative proximity within the same chromatin domain could be advantageous – for example by reducing the search space of the enhancer for the promoter (Williamson et al., 2011; Benabdallah and Bickmore, 2015). Consistent with this model, we have previously shown that, in the limb, expression levels of a reporter gene inserted into several positions across the whole *Shh* regulatory domain, is highest when the reporter inserts close to either the ZRS or *Shh* compared to insertion sites within the intervening gene desert (Anderson et al. 2014). These data suggest that ZRS-induced expression requires direct or indirect interactions with the target gene and these interactions are optimised by minimising the search space within a constrained chromatin domain. Whether the actual co-localisation of the ZRS with *Shh* in the ZPA is a cause or consequence of limb-specific Shh activation remains to be determined.

Acknowledgements

We thank the staff of the IGMM imaging facility and technical services for their assistance with imaging and sequencing. We thank the Dostie lab at McGill University for access to and the use of their Torrent 5C data transformation pipeline on the McGill University local galaxy server. The Super-resolution imaging experiments were conducted using the facilities provided by the Edinburgh Super-Resolution Imaging Consortium (ESRIC). This work was supported by the Medical Research Council, UK.

Author Contributions

IW and LAL conducted the experiments, and contributed to both the experimental design and writing of the paper. REH and WAB contributed to the design of the project and the writing of the paper

References

Amano, T., Sagai, T., Tanabe, H., Mizushina, Y., Nakazawa, H., Shiroishi, T., 2009. Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. *Developmental cell*. 16(1), pp. 47-57.

Anderson, E., Devenney, P. S., Hill, R. E., Lettice, L. A., 2014. Mapping the Shh long-range regulatory domain. *Development*. 141, pp. 3934-3943.

Belmont, A. 2014. Large-scale chromatin organisation: the good, the surprising , and the still perplexing. *Curr. Opin Cell Biology*. 26, pp. 69-78.

Benabdallah, N. S., Bickmore, W. A., 2015. Regulatory domains and their mechanisms. *Cold Spring Harb. Symp. Quant. Biol.* pii: 027268.

Berlivet, S., Paquette, D., Dumouchel, A., Langlais, D., Dostie, J., Kmita, M., 2013. Clustering of Tissue-Specific Sub-TADs Accompanies the Regulation of HoxA Genes in Developing Limbs. *PLoS Genetics*, 9(12).

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., Ren, B., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 485, pp. 376-380.

Dostie, J., Dekker, J., 2007. Mapping networks of physical interactions between genomic elements using 5C technology. *Nat. Protocols*. 2, pp. 988-1002.

Dostie, J., Zhan, Y., Dekker, J., 2007. Chromosome conformation capture carbon copy technology. *Curr. Protoc. Mol. Biol.* Chapter 21: Unit 21 14.

Ferraiuolo, M. A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X. Q. D., Nadler, M., Blanchette, M., Dostie, J., 2010. The three-dimensional architecture of Hox cluster silencing. *Nucl. Acids Res*. 38, pp. 7472-7484.

Fraser, J., Ethier, S. D., Miura, H., Dostie, J., 2012. A torrent of data: mapping chromatin organization using 5C and high-throughput sequencing. *Methods Enzymol*. 513, pp. 113-141

Fraser, J., Rousseau, M., Blanchette, M., Dostie, J., 2010. Computing chromosome conformation. *Methods Mol. Biol.* 674, pp. 251-268

Fraser, J., Williamson, I., Bickmore, W. A., Dostie, J., 2015. An Overview of Genome Organization and How We Got There: from FISH to Hi-C. *Microbiology and Molecular Biology Reviews*. 79(3). pp. 347-372.

Ghavi-Helm, Y., Klein, F. A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., Furlong, E. M., 2014. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature*. 512, pp. 96-100.

Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M., Epstein, D. J., 2006. A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development*. 133. pp. 761-772.

Klopocki, E., Ott, C-E., Benatar, N., Ullmann, R., Mundlos, S., Lehmann, K., 2008. A microduplication of the long range SHH limb regulator (ZRS) is associated with triphalangeal thumb- polysyndactyly syndrome. *J. Med. Genet.* 45, pp. 370-375

Lajoie, B. R., van Berkum, N. L., Sanyal, A., Dekker, J., 2009. My5C: web tools for chromosome conformation capture studies. *Nat. Methods*. 6, pp. 690-691

Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E., de Graaf, E., 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Human Molecular Genetics*. 12(14). pp. 1725-1735.

Lettice, L. A., Hill, A. E., Devenney, P. S., Hill, R. E., 2008. Point mutations in a distant sonic hedgehog cis-regulator generate a variable regulatory output responsible for preaxial polydactyly. *Human Molecular Genetics*. 17(7). pp. 978-985.

Lettice, L. A., Horikoshib, T., Heaney, S. J., van Barenb, M. J., van der Lindee, H. C., Breedvelde, G. J., Joossee, M., Akarsuf, N., Oostrae, B. A., Endod, N., Shibatag, M., Suzukih, M., Takahashih, E., Shinkai, T., Nakahorii, Y., Ayusawaj, D., Nakabayashik, K., Schererk, S. W., Heutinke, P., Hill, R. E., Nojic, S., 2002. Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *PNAS*, 99(11). pp. 7548-7553.

Lettice, L. A., Williamson, I., Devenney, P. S., Kilanowski, F., Dorin J., Hill, R. E. 2014. Development of five digits is controlled by a bipartite long-range cis-regulator. *Development*. 141(8). pp. 1715-25.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Duboule, D., 2011. A Regulatory Archipelago Controls Hox Genes Transcription in Digits. *Cell*. 147, pp. 1132-1145.

Morey, C., Da Silva, N. R., Perry, P., Bickmore, W. A., 2007. Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development*. 134, pp. 909-19.

Nora, E.P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N. L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blu, N., Dekker, J., Heard, E., 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485, pp. 381-385.

Patel, N.S., Rhinn, M., Semprich, C. I., Halley, P. A., Dolle, P., Bickmore, W. A., Storey, K. G., 2013. FGF signalling regulates chromatin organisation during neural differentiation via mechanisms that can be uncoupled from transcription. *PLoS genetics*, 9(7), p.e1003614.

Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., Shiroishi, T., 2005. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development*. 132(4), pp. 797-803.

Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, T., Noda, T., Shiroishi, T., 2004. Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mamm. Genome*. 15 (1), pp. 23-34.

Sun, M., Ma, F., Zeng, X., Liu, Q., Zhao, X-L., Wu, F-X., Wu, G-P., Zhang, Z-F., Gu, B., Zhao, Y-F., Tian, S-H., Lin, B., Kong, X-Y., Zhang, X-L., Yang, W., Lo, W. H-Y., Zhang, X., 2008. Triphalangeal thumb-polysyndactyly syndrome and syndactyly type IV are caused by genomic duplications involving the long range, limb-specific SHH enhancer. *J. Med. Genet.* 45, pp. 589-595

Toomre, D. & Bewersdorf, J., 2010. A new wave of cellular imaging. *Annual review of cell and developmental biology*. 26, pp. 285-314.

Wieczorek, D., Pawlik, B., Li, Y., Akarsu, N. A., Caliebe, A., May, K. J. W., Schweiger, B., Vargas, F. R., Balci, S., Gillessen-Kaesbach, G., Wollnik, B., 2010. A specific mutation in the distant sonic hedgehog (SHH) cis-regulator (ZRS) causes Werner mesomelic syndrome (WMS) while complete ZRS duplications underlie Haas type polysyndactyly and preaxial polydactyly (PPD) with or without triphalangeal thumb. *Hum. Mutat.* 31, pp. 81-89

Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R. S., Paquette, D., Dostie, J., Bickmore, W. A., 2014. Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes & development*. 28, pp. 2778-2791.

Williamson, I., Eskeland, R., Lettice, L. A., Hill, A. E., Boyle, S., Grimes, G. R., Hill, R. E., Bickmore, W. A., 2012. Anterior-posterior differences in HoxD chromatin topology in limb development. *Development*. 139(17), pp. 3157-67.

Williamson, I., Hill, R.E. & Bickmore, W. a., 2011. Enhancers: From Developmental Genetics to the Genetics of Common Human Disease. *Developmental Cell*. 21(1), pp. 17-19. Available at: http://dx.doi.org/10.1016/j.devcel.2011.06.008.

Figures

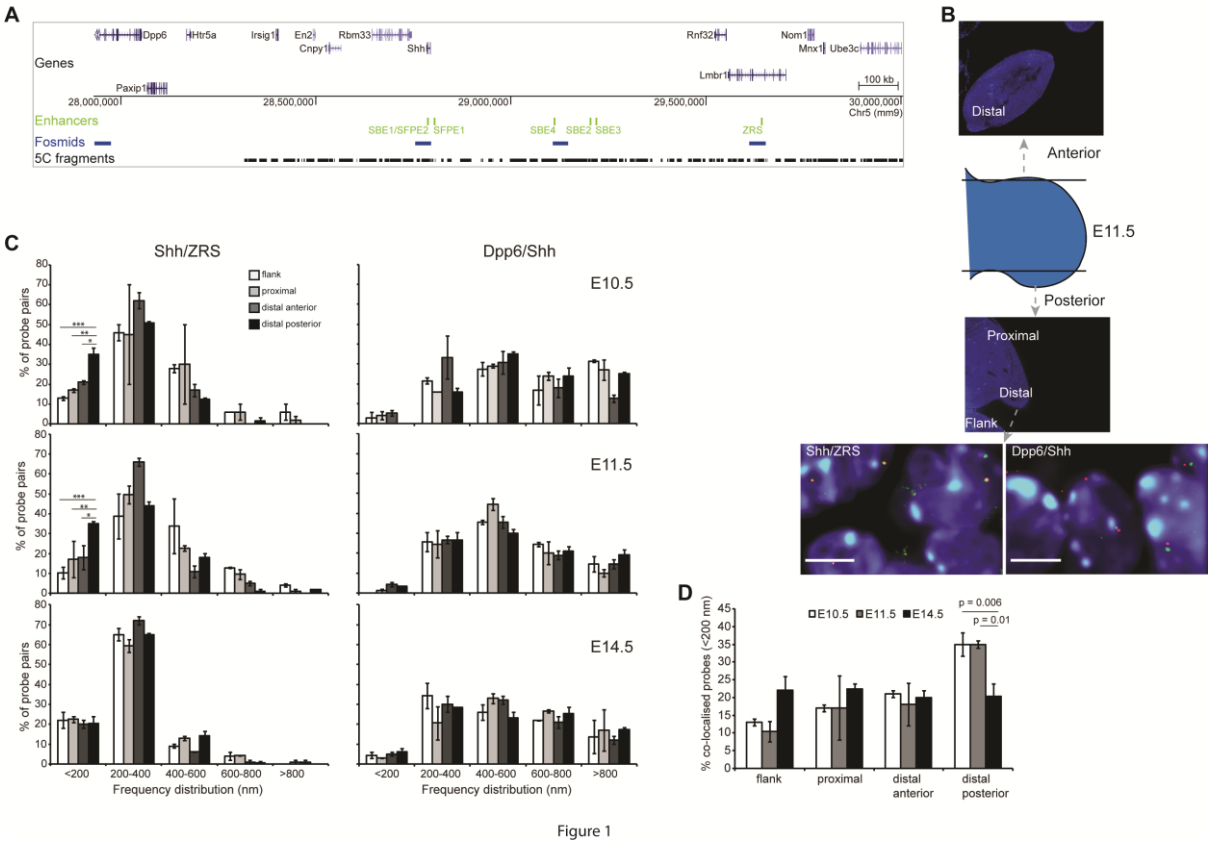

Figure 1

Figure 1. ZRS-*Shh* proximity in the ZPA at E10.5 and E11.5. (**A**) (Top) Location of genes over a 2 Mb murine genomic locus containing *Shh*, with the position of tissue-specific *Shh* enhancers shown below in green. The bottom two tracks show the locations to which the fosmid probes used for FISH hybridize (blue) and the 3C fragments amplified for 5C (black). (**B**) Schematic indicating the position and plane of the tissue sections taken through the anterior and posterior parts of the E11.5 forelimb bud. Distal and proximal parts of the posterior limb bud and the distal anterior limb bud are shown, as is the flank mesoderm. Below are images of nuclei from E11.5 ZPA tissue sections showing *Shh*/ZRS and *Shh*/*Dpp6* probe pairs. Scale bars = 5μm. (**C**) Frequency distributions of FISH inter-probe distances (d) in 200 nm bins, between *Shh* and ZRS (left column), or *Shh* and *Dpp6* probes (right column) in proximal and distal (anterior and posterior) regions of the murine forelimb bud and adjacent flank at E10.5, E11.5 and E14.5 (n = 70-130 alleles). For E10.5 and E11.5 sections distal posterior limb = ZPA. Error bars represent SEM obtained from two or three different

tissue sections from 1-2 embryos. The statistical significance between data sets was examined by Fisher's Exact Tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (**D**) Comparison of the proportion of co-localised *Shh*/ZRS probe pairs (<200nm) across the three temporal developmental stages for distal anterior and posterior and proximal forelimb tissue and flank tissue. Error bars represent SEM obtained from two or three different tissue sections. The statistical significance between data sets was examined by Fisher's Exact Tests.

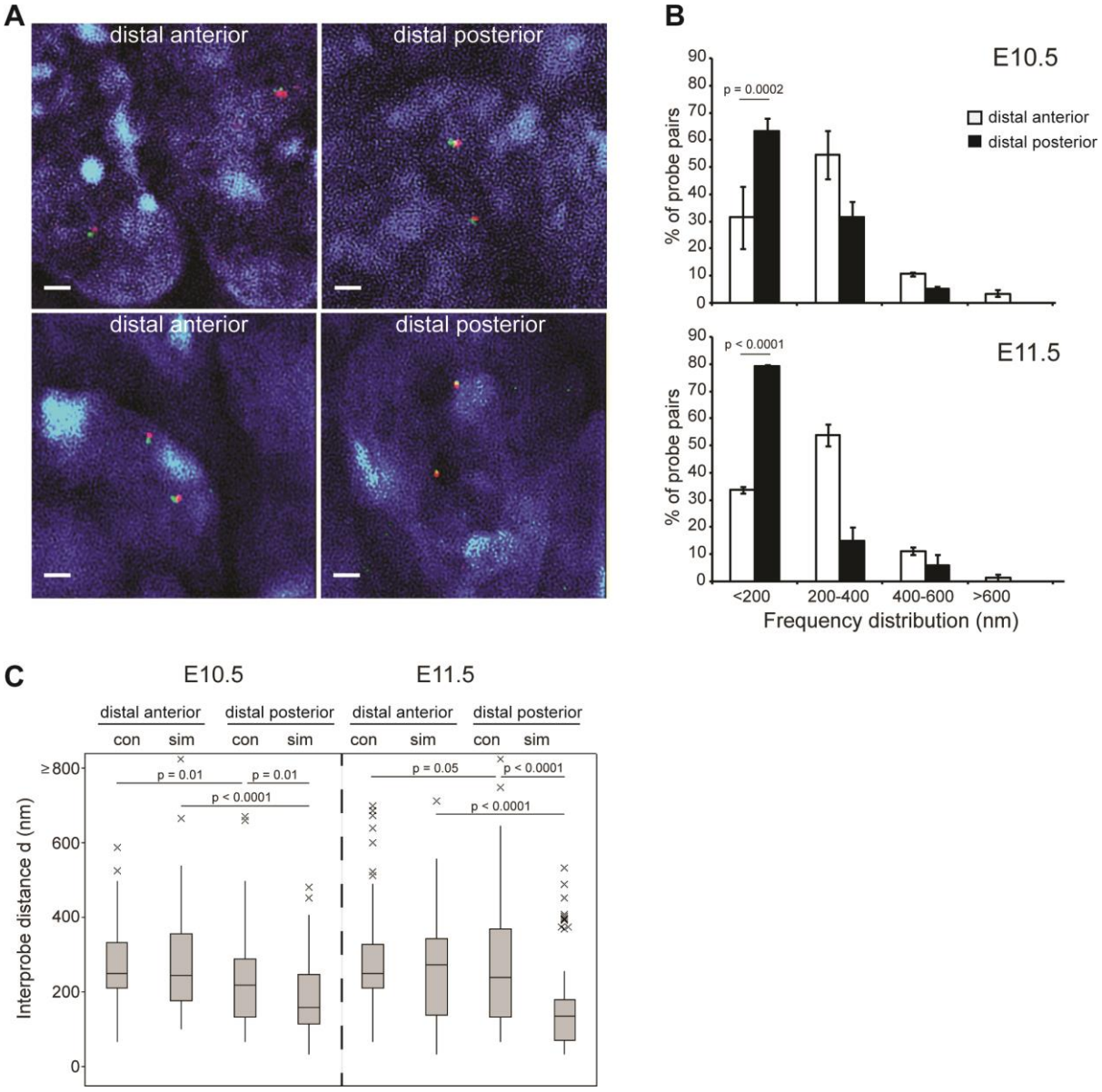

Figure 2. Super-resolution imaging identifies the majority of *Shh*-ZRS probes as co-localised in ZPA tissue. (**A**) Nuclei captured by super-resolution SIM imaging from the distal forelimb of E10.5 and E11.5 embryos after FISH with *Shh* and ZRS probe pairs. Scale bars = 1 μm. (**B**) Frequency distributions of *Shh*-ZRS inter-probe distances (d) measured from SIM images in 200 nm bins, in distal anterior and distal posterior regions of the murine forelimb at E10.5 and E11.5. n = 67-100 (alleles). Error bars represent SEM obtained from two different tissue sections from 1 embryo. The statistical significance between data sets was examined by Fisher's Exact Tests. (**C**) Boxplots show the distribution of *Shh*-ZRS inter-probe distances (d in nm) in E10.5 and E11.5 distal anterior and distal posterior captured by

conventional (con) and structured illumination (sim) microscopy. Line = median, box = interquartile range, whiskers = 95% range. The statistical significance between data sets was examined by Mann-Whitney U Tests.

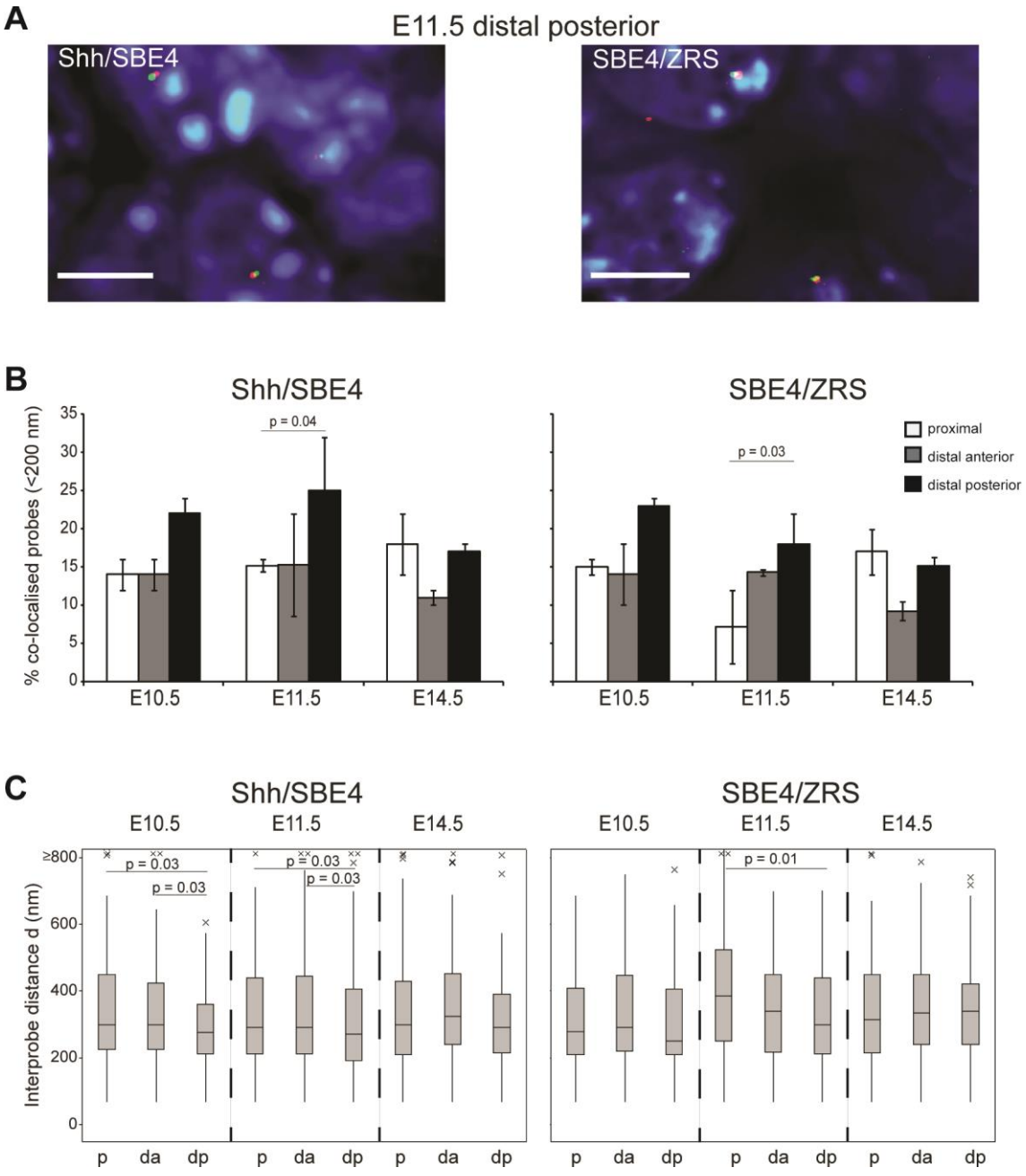

Figure 3. The *Shh*-ZRS regulatory domain is maintained in a compact chromatin conformation in expressing and non-expressing tissue. (**A**) Images of representative nuclei from E11.5 ZPA tissue showing FISH signals for *Shh*/SBE4, SBE4/ZRS probe pairs. Scale bars = 5 μm. (**B**) Comparison of the proportion of co-localised *Shh*/SBE4 and SBE4/ZRS probe pairs (<200nm) across the three temporal developmental stages for proximal and distal anterior and posterior (ZPA in E10.5 and E11.5 sections) forelimb tissue (n = 70 – 100 alleles). Error bars represent SEM obtained from two or three different tissue sections from 1-2 embryos. The statistical significance between data sets was examined by Fisher's Exact

Tests. (**C**) Boxplots showing the distribution of interprobe distances (d) in nanometres between *Shh*/SBE4 and SBE4/ZRS in E10.5, E11.5 and E14.5 proximal (p), distal anterior (da) and posterior (dp) forelimb. The statistical significance between data sets was examined by Mann-Whitney U Tests.

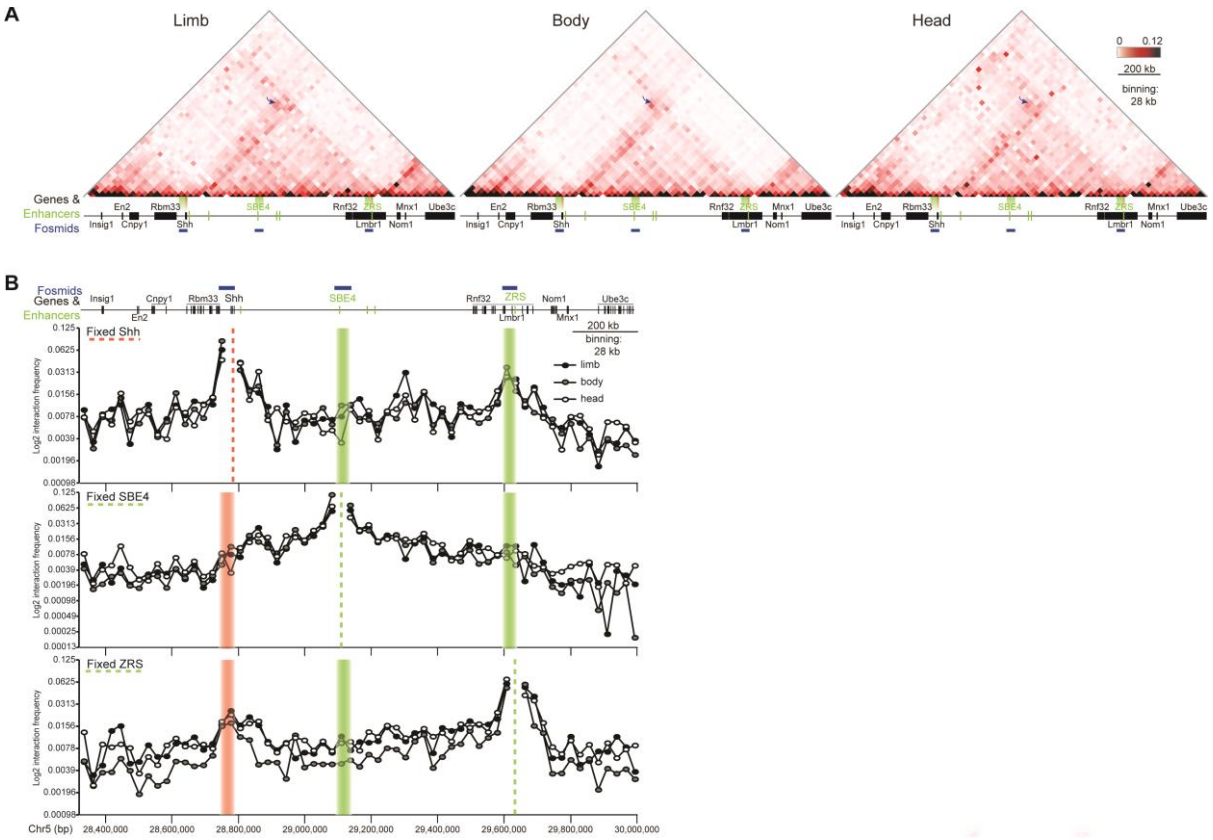

Figure 4. 5C-seq identifies enriched interactions between *Shh* and ZRS in E11.5 embryos. (**A**) Heat- maps showing 5C data from cells of the limbs, bodies and heads of E11.5 embryos, across the 1.7-Mb Shh region shown in Figure 1. Heat map intensities represent the average of interaction frequency for each window, colour-coded according to the scale shown. Interaction frequencies were normalized based on the total number of sequence reads in the 5C data set and the data shown is binned over 28-kb windows. Arrows indicate interaction frequencies between windows containing *Shh* and ZRS. Data for biological replicates are in Supplemental Figure S3A and unprocessed normalized data are shown in Supplemental Figure S4. (**B**) Virtual 4C analysis obtained by extracting 5C interactions with viewpoints fixed at *Shh*, SBE4 and ZRS. Dashed lines indicate the position of the fixed viewpoint from the *Shh* genomic region (orange) or regulatory elements (green). Data from limbs are in closed black circles, bodies is closed grey circles and heads in open circles. Genome coordinates on Chr5 are from the mm9 assembly of the mouse genome.

Supplemental Tables and Figures

Table S1. Fosmid Probes

Region	Whitehead (Sanger) Name	Ensemble name	Coordinates Start	End	Size (bp)
Dpp6	WI1-2157A11	G135P600264D6	27932527	27975636	43109
Shh	WI1-482L15	G135P64333A4	28754458	28795879	41421
SBE4	WI1-469P2	G135P600205H10	29107140	29147593	40453
ZRS	WI1-121N10	G135P600929F6	29611727	29653695	41968

Names are Ensembl (r 45) (http://jun2007.archive.ensembl.org/Mus_musculus/index.html). Mouse genome assembly number: NCBI m37

Table S2. Co-localisation frequency (<200 nm) of *Shh* and ZRS probes in distal and proximal limb and adjacent flank tissue of E10.5, E11.5 and E14.5 embryos at normal resolution (*x* & *y* bins = 67 nm, *z* steps = 200 nm)

	E10.5	E11.5	E14.5
Tissue	**Colocalisation frequency (%)**		
Distal posterior	35	35	20
Distal anterior	21 (p = 0.02)	18 (p = 0.01)	20
Proximal	17 (p = 0.003)	17 (p = 0.005)	22
Flank	13 (p = 0.0002)	10 (p = 0.0002)	22

Statistical analysis of data for Fig. 1C. p-values from Fisher's Exact Tests.

Table S3. Co-localisation frequency (<200 nm) of *Shh* and ZRS probes in distal anterior and posterior tissue of E10.5 and E11.5 embryos at super resolution (*x* & *y* bins = 32 nm, *z* steps = 120 nm)

	E10.5	E11.5
Tissue	**Colocalisation frequency (%)**	
Distal posterior	63	79
Distal anterior	31 (p = 0.0002)	34 (p < 0.0001)

Statistical analysis of data for Fig. 2B. p-values from Fisher's Exact Tests.

Table S4. Median interprobe distances for *Shh* and ZRS probes in distal anterior and posterior tissue of E10.5 and E11.5 embryos at super resolution

	E10.5	E11.5
Tissue	**Interprobe distance (nm)**	
Distal posterior	160	136
Distal anterior	246 (p < 0.0001)	275 (p < 0.0001)

Statistical analysis of data for Fig. 2C. p-values from Mann-Whitney U Tests.

Table S5. Median interprobe distances for *Shh-Dpp6*, *Shh*-SBE4, SBE4-ZRS and *Shh*-ZRS probes in distal and proximal limb and adjacent flank tissue of E10.5, E11.5 and E14.5 embryos

Tissue	*Shh-Dpp6*	*Shh*-SBE4	SBE4-ZRS	*Shh*-ZRS
E10.5	**Interprobe distance (nm)**			
Distal posterior	593	276	250	221
Distal anterior	493 (p = 0.005)	300 (p = 0.03)	291	250 (p = 0.01)
Proximal	611	300 (p = 0.03)	280	263 (p = 0.0001)
Flank	571	314 (p = 0.02)	250	291 (p < 0.0001)
E11.5	**Interprobe distance (nm)**			
Distal posterior	523	272	300	241
Distal anterior	480	291 (p = 0.03)	341	250 (p = 0.05)
Proximal	471	291 (p = 0.03)	385 (p = 0.01)	250 (p = 0.0004)
Flank	512	341 (p = 0.02)	345	406 (p < 0.0001)
E14.5	**Interprobe distance (nm)**			
Distal posterior	540	291	341	250
Distal anterior	479	324	334	241
Proximal	549	300	314	250
Flank	474	295	287	241

Statistical analysis of data for Fig. S1B. Interprobe distances are median values, p-values from Mann-Whitney U Tests.

Table S6. Mouse 5C primers for *Shh* and *USP22* regions

Fragment	Type	Genomic sequence (5' to 3')	HindIII position Start	HindIII position End
Shh (chr.5; mm9)				
3	R	CTTCCAGTATTGGGTTACAGTTAATGGAGT	28317087	28319149
5	F	AGCATAGAGTGTGTGTAGGTGCTGCCTAAG	28319935	28324734
7	R	CTTTCTCTCATCCCTACACTAACCAGGCCT	28325862	28331585
8	F	GGTAAGAGTCCCAAAGAACAGCTTGTTAAG	28331586	28333893
9	R	CTTGACACACCTACCCTCTAAGTAATCAAT	28333894	28336649
10	F	AACACCTCTAGCATGATAGCACTTTGCAAG	28336650	28340960
11	R	CTTAGGATGTGCCTCTACTGTGGGGG	28340961	28344675
12	F	CAAAACCCTAGAAGCCACAGGGACCAAG	28344676	28347365
13	R	CTTGCCAGTTTATCTAGGTAGCCTGCCAG	28347366	28348911
18	F	GGAGACCCACACTAAGGGCCTCAAG	28355883	28370234
19	R	CTTGGATTGGCGTGGCTCTGAGTCAT	28370235	28373716
20	F	TGTGTTTTAGGGATGAGGGATTCTTTAAAG	28373717	28374948
25	R	CTTCCTTTCTGGTATCTATTGACCTTCCCT	28385728	28388130
27	F	TTTGATAGTGCTGTTTCCTGTGGCTAGAAG	28388373	28403099
30	R	CTTCTCCTGTAAGATGGCAATTTATTTATT	28416843	28419608
31	F	TTCTAAATATAATCCAGAGAGAAGGCTAAG	28419609	28428926
32	R	CTTTGTCATTCCACCATGGTTGTGGTGAAC	28428927	28447705
34	F	CAATAAAGGTAGAACTTGGGCCCAGTGAAG	28452395	28454456

35	R	CTTGAGCTCACATATGGGACACTCTTGACA	28454457	28458325
38	F	ATGGGGCCGGATTTAACTCAACAATCAAAG	28463308	28463573
41	R	CTTCTCTAGCTAGGCCAGCATAATGTACCG	28469222	28469868
42	F	AAAAATAAAAAGGAGGCCTGGATTCTGAAG	28469869	28470959
44	R	CTTAGCTCAAAATGTAGGAAATGGCCATTC	28476584	28484052
45	F	AAAATCTCCCTGGAGTCAAAGGGTTAAAAG	28484053	28486216
46	R	CTTGTGAACAGTCCCACCAGGTCACTG	28486217	28491939
47	F	CCACCCCCAGTATCTGCAACCTCAAG	28491940	28496579
49	R	CTTAAGGTGGGGGTGACACAGTCCAAAG	28504618	28506235
50	F	GCAAGAGCCCACCAGGGTCAAG	28506236	28509301
52	R	CTTCAGTTTGGGTGACACATGCAGGACAAA	28509323	28510999
53	F	GTGATGTCTCCCCTGTGAGCAGGAAG	28511000	28511814
56	R	CTTATCCTCTTCTGTGTCTAGTTGAAGTGG	28511842	28514356
57	F	TTCTCATGTCAAGATCCACATAATATCAAG	28514357	28515715
58	R	CTTGTTGAGCTCATGTAAATGCCTATGGAT	28515716	28528247
59	F	TTACATATCAGCTGCTGTATCCATCACAAG	28528248	28532168
61	R	CTTGAAAAGCACAGATAAAAATGCCATTTG	28534318	28537153
62	F	CACAGGGCTCTTTCATAGCCTAAGAACAAG	28537154	28541461
63	R	CTTCCTTGGGCAAGTGGTATCTTCCTTAGC	28541462	28542832
64	F	CACAGTCACCCCCTGTACCCCAAAG	28542833	28543234
65	R	CTTTGACAAAGTGATGCATCTAAGATCCTA	28543235	28543365
68	F	GGTACTTCTACAGTGGGGGAGGATGTAAAG	28547549	28547729
69	R	CTTCTCATATGACTCTGGTTTCTTGGCCCA	28547730	28548357
73	F	TCTATGTATAAGCCACACCAAGGAAAGAAG	28559901	28566954
74	R	CTTAGCATGGGACTCAGAAAACAAAATAGG	28566955	28570463
75	F	TGGCAAATCAGAAAAACTCTTTGGATTAAG	28570464	28571000
76	R	CTTGAAATTGAAGTATCTCTCTCAGCACCT	28571001	28572084
77	F	GAGTTCAAGAGCCCCCAAATCCCTCTAAG	28572085	28576299
78	R	CTTCCAGAAGATCTGCAGCAACTCTCTCTC	28576300	28577015
79	F	GGAGGCAGAGCCTCTGAGTCACAAG	28577016	28580658
80	R	CTTGAAGCATGTGTGGACTCCCATTCTTCT	28580659	28581218
81	F	CTCCAGACTGAGACCTTCCTGAGACCAAAG	28581219	28585322
83	R	CTTGCGGTTGTAGCTAAGAGTGAATTTGAA	28585390	28593411
84	F	GGAAAGGTACTCTGGGGTGCATCACAAAG	28593412	28595934
86	R	CTTTGGGGTGGGAACAAGGAGACTTCAC	28604862	28605362
90	F	GACCTGACACCTTTGGGATGAAAGTGTAAG	28611343	28619284
92	R	CTTGACAGAGGAGCCTAAAAGGTGACTTAA	28619329	28620319
93	F	ATTGGCTATGTAGATGAAGATGGTCCCAAG	28620320	28627997
94	R	CTTTCATTGAGACACTCCCTCAGCCTCAGT	28627998	28631580
95	F	CGTGTTACAGTTAGCTACACCCTCAAAAAG	28631581	28635856
96	R	CTTAGCAGTTTCTGTAAAAAATAAAAGTAC	28635857	28638808
97	F	GTAGCGTTCGCGCGCCTCAAG	28638809	28643392
98	R	CTTGTGCAGTACTAAATCATAATGCCATAA	28643393	28646446
99	F	TAGCATAGGGGTTATGGATGGACTCAAAAG	28646447	28647705
100	R	CTTTAAAAGGTACAATGATAGAAGAAATAG	28647706	28650692
101	F	TGAGTTTTCAGTAACCACTATAAAAAGAAG	28650693	28651422

102	R	CTTCCTATGGCTGAGAACTGCTTAGATAAT	28651423	28652423
103	F	GCTGATTCCTTTGCTGACTGGAGTGTGAAG	28652424	28656081
106	R	CTTTTGGTTGTTTCAACCATTTTTCACTTA	28662072	28668894
107	F	GTCTTCTTATCCCTGGGTAAATTGTTAAAG	28668895	28669869
109	R	CTTCCTCTCTCAGGAAACCAGTCTTCTGAG	28676542	28681826
111	F	CAGCATGGCTGTGAGGGAAAGTAGCTAAG	28681874	28682959
112	R	CTTCCAACAGTACATTATCCTAAGCGTCTA	28682960	28684009
114	F	CAATTTCAGTGCCAGCCTCTCTCGGTAAG	28685148	28687114
115	R	CTTACACCAATTAACTTCTTAGAAGTAGAC	28687115	28688407
117	F	CCTGTTTGCACTGTGTCTTCTCACAGGAAG	28689808	28693301
118	R	CTTAAATCACGAAGTACTGAGGCTTACCAA	28693302	28698428
119	F	AAAAGAAAAAGAATCACCATGACTCTTAAG	28698429	28705774
120	R	CTTATTCACAGGCCATTCTGGTAGGAACAT	28705775	28710085
121	F	CATGGTAACATCGTGTGTAGATAGAAAAAG	28710086	28710865
122	R	CTTATGTGGAGAACTAACACCATCATAAAT	28710866	28715944
123	F	AACTTTGTGGCACCTTTCTCCTCCAGAAAG	28715945	28717043
124	R	CTTTACCACTGCGGAAGGGGGAAAACA	28717044	28718937
128	F	CCTTCTGTGTGATCATCTGACACATAAAAG	28723250	28724961
129	R	CTTATACTGGGTGGAGGTCAATTCTGGACT	28724962	28731808
130	F	GTTGCAGACTGAGGGGCTCCAGAAG	28731809	28735435
131	R	CTTTAACCTGGCTCTGCTCTCAGAATGAGG	28735436	28738817
132	F	ATTTTCTTGTGGCATTATTAGGCAGGTAAG	28738818	28743466
133	R	CTTTAATGTTCTGGTTTGTTGTTGTTCTAA	28743467	28748045
134	F	CCTGCAGTCAGGGAACCGAGAGAAG	28748046	28750565
139	R	CTTTTATGTGGCTCTGCTTTTGTATTACAA	28759100	28763978
140	F	ATATTTGGATGTTCTGTCAGTGGCCTGAAG	28763979	28767078
143	R	CTTGTTCCCCGTACCCACATAAAAGGCC	28770697	28780563
144	F	CCAGAGACCCCTCCATCTGCTCAAAG	28780564	28782634
145	R	CTTTTCCCTCACCCCATTGAAAGAAGGGAG	28782635	28789524
147	F	CATCTGATTGGCCAAGCCGCACAAG	28789547	28793709
149	R	CTTACTACCAGTCCTTTGCTCTGTCTAATA	28794058	28797606
151	R	CTTGGCTAACATTGGACAACCCAAGTGTTT	28798546	28799298
152	F	TAGAAAAGATGCTGGGAACCTCATTCTAAG	28799299	28801546
154	R	CTTCTGGACACCCAGAAATGTGCGTCTC	28805019	28805807
155	F	GTGGAGCCATCATGGAAATTGCATGGAAG	28805808	28806894
158	R	CTTCTCCTAAGAACAGCTAGACCTATGCAC	28821671	28827662
162	F	AGAACCACAGGATACCCATAAGAGCCAAAG	28837190	28843289
163	R	CTTCAAAGCTGCAGTGCTTTGAAGTGTCTG	28843290	28847606
168	F	ATATCTACACCAGCTTTCTAAAATGGAAAG	28870608	28876511
169	R	CTTTATTGCCAGGTCAAATGATTTAAACAT	28876512	28878134
170	F	CTGTGATCTGAAGGTGTAAGCTGAGATAAG	28878135	28878784
171	R	CTTTGAAGGAGACCCTATTTCCTATGTGGG	28878785	28879286
173	F	GGTGTGCAGCCAGTGTGCATATTAGACAAG	28879305	28880400
174	R	CTTTATTCACCTCTGACATGCAAGCCAACA	28880401	28880839
176	F	GGCTGCAAAAGTTGGGTCTCATTTGTGAAG	28882358	28883821
177	R	CTTTGGAAGGCTGGGTGGTCAGC	28883822	28884803

178	F	CTTGGGTAAAGCTATACTGGATGCGGCAAG	28884804	28887360
179	R	CTTCCAAAATTCTATTTTGGGAAAAAATGA	28887361	28889232
180	F	TTACCTAGGTAACTGCTGCCCTTTCAGAAG	28889233	28895175
181	R	CTTCCAGAAACTGTTTACTTCCTTCTGGAG	28895176	28899338
191	F	TTGTTATTCCTGGGAGACTTAATTGGCAAG	28923745	28923931
193	R	CTTCCAGGTCACGTTAAAGATATTTCAGTA	28925837	28929215
197	F	CCACAACCTGCCCCTACGGTGTAAAG	28942022	28945936
198	R	CTTAGAAGGACCCTAGAATGGTCCCCTGAA	28945937	28949569
199	F	GGTATGGGATATCCTTTGGGGTTCACTAAG	28949570	28950647
200	R	CTTTGAATGAGGTCAATAAAAATCTACCTC	28950648	28951388
201	F	CTTCACTCCCTCAGTTACCAAGCCACAAAG	28951389	28952477
204	F	CAGGATGCCTTAGGAGACACGAGGAAG	28954090	28954407
207	R	CTTTTCAAAAAGCCACATGAAAATCTACTA	28971638	28978794
212	F	CAGGCACCTAAGTGTTAGAGAAGTTGGAAG	28997868	29000514
213	R	CTTATGAGTGCAAGGTCTGCCCAGGTTG	29000515	29003765
214	F	TACAAGTCTCATCTGAGCCCTCCAAAAAAG	29003766	29004021
215	R	CTTAGTTGTTCCTTCAGTGTCAGTTACTTC	29004022	29005070
216	F	GATGGGTCTTTCAGAGTCTGTTCCCTGAAG	29005071	29006498
217	R	CTTTTATTAAAAGGCCATGGGGCCATGGAG	29006499	29008669
218	F	TGTTCACCAAAATTTATTCTAAAAGGCAAG	29008670	29011994
219	R	CTTCTAATGAACCTGCTCCTGACCGCATG	29011995	29013866
220	F	CAAGATTACCCTGAAGTGCCGGTACAGAAG	29013867	29014066
221	R	CTTCAGGAGTCTAACTGCAATAATAATGCA	29014067	29020057
222	F	GCTCTGTGTGTCCTTCAGCTCTCTGAAG	29020058	29022523
223	R	CTTTGCTGTTGGTCAAAGGTAGCAGCTGA	29022524	29028120
224	F	AGCAGGCTTCCTCCTAGGATTATAATGAAG	29028121	29029629
225	R	CTTCGGGCCCCGGAGGGAGG	29029630	29032468
226	F	CTGAGTCTCAAGCAGCTAGCTTTCAGAAAG	29032469	29038542
227	R	CTTTTGAGAGGTGTGAGACAATCAAAATAA	29038543	29039555
230	F	CTTCCCAGGCTTTGAAGGGAACACACTAAG	29048476	29050003
231	R	CTTCAATTTGTGAGCCTCTACAAAACCTCA	29050004	29051239
232	F	GGTCCACTGGCAGCCCAAGAAG	29051240	29058423
233	R	CTTCCTTTTCATCTTGAATCAGCCTATAGA	29058424	29060961
234	F	AGTGTTTAGGGTTCTAAGGACATGGCAAAG	29060962	29068732
237	R	CTTTCTCCCAGTGATGCTGTTAGTTGTTCC	29078836	29081207
238	F	AGTTTCCTTGAAAATCATGGCCCAGCAAAG	29081208	29082327
239	R	CTTGCAATACTTCCAAGAGAAAGAGAACAC	29082328	29083134
241	F	TCACACACCTGGCAGCTGGACAAG	29084039	29096781
242	R	CTTTCAATTACTGCTCTAGCACAATGCCTG	29096782	29100883
243	F	AAGATAAAAACCTTCCATCTTAGAACAAAG	29100884	29101246
244	R	CTTGAAAAGAGGAATATTCGGGAATGGAGG	29101247	29109459
245	F	TAGGAAAAGGGACTTAATTTGTCATCAAAG	29109460	29111881
246	R	CTTAGATCTGGTGGTAGAAGGGATACTGGA	29111882	29113252
247	F	CGTATAGACTTAAATGAATTAGAACAAAAG	29113253	29117196
248	R	CTTTGCTGTCTGCTCAGAGGAGAACTCTGT	29117197	29119739
251	F	AGAAGGATTGTTTTATTCCTGTCCTTAAAG	29130489	29131877

252	R	CTTAGGCCATGAAGGAAGATGGCTTTGACA	29131878	29136802
254	F	GTTTTTTTCTGGCACAAGCTACCACTTAAG	29137043	29154044
255	R	CTTCCAATGATCCTCTCTGCTCCAATGAAA	29154045	29155707
256	F	TTTATTCTAACACTTATCCCATCCTGCAAG	29155708	29159386
258	R	CTTGGGGACCATTTAAAATATGTTCTAGAT	29159409	29160128
259	F	CTATATTTATTTAAAGTACAAAAACCTAAG	29160129	29166826
260	R	CTTGCCATGTAAGATGGGATATCTGGCCAG	29166827	29168466
261	F	GCACCGAGACCAGCTTCTGAGATCAAG	29168467	29173627
262	R	CTTATTTATCAAGTACAGTTGCTCAAGTAT	29173628	29173865
263	F	CATTAAACTGTGTTGAACCTATTTATTAAG	29173866	29174415
264	R	CTTGCAGCTCCAGACATTTACCCATCTCTG	29174416	29179395
265	F	TGGCACCTCAAATTGAGACCTTGCTTTAAG	29179396	29180591
266	R	CTTCCTCCTTACCCCGGCTGCTAATC	29180592	29184579
267	F	AGTGTTTCTATTTTTATGCTGAGTCCCAAG	29184580	29185047
268	R	CTTTGATATTACAGTGATGAATTGATATGT	29185048	29185852
269	F	TGAGAACAGTATTTTACTTAATTTGAGAAG	29185853	29188276
271	R	CTTAGACTTCCTTTTAGAACATTCTTGGTT	29188338	29191726
272	F	TTGAAAGCTGATTTCAAACAATGATTAAAG	29191727	29196352
274	R	CTTATCAACAACACCTGCATTTTAAAAGAC	29198345	29199536
275	F	AGCTATCATTTGGTTAAAAACTGTTAGAAG	29199537	29203806
276	R	CTTCATGCTGGCAGACAAAGTAAATTCGGG	29203807	29203920
278	F	GGCACATGCTGGGTCCCAGATAAG	29206157	29209241
279	R	CTTCCTTGTAAATATCTGGAATAGAAAGAG	29209242	29216468
280	F	CATGGGAGGTGTCAAACGGATTGGTGAAG	29216469	29220327
281	R	CTTGGAGATTTCCAGGTCTGCTTCTACATT	29220328	29223230
283	F	GCCACAGTCTGGAAGCACAGATCCAAG	29230618	29233018
284	R	CTTCCATCAGATCTGTGTTGTCCAAGAATC	29233019	29236380
285	F	CCAGGGTCATCTATGTGCATGCTCACAAG	29236381	29237108
287	R	CTTGCTAACACCGTTTGAGGTGGAGATGC	29237129	29240127
288	F	TGTCATCCCTTGTTCTAGTTTGACTCTAAG	29240128	29243266
289	R	CTTCTCCTGGTCATCATCTGGCCAGTATCA	29243267	29246728
290	F	GCGTAGTAAACACGGGGGTTATAAGCTAAG	29246729	29247568
291	R	CTTGGCTACACTGCATTTCCTTGTATTATA	29247569	29250429
292	F	GGGAGCAATTCTTAAGAGTGCTTTTCTAAG	29250430	29250736
295	R	CTTTGAAGATTTCCAAGAGTTCAAACCCAG	29253038	29259186
296	F	CTAGGTTTCTTGGGGAGAAGGGCTATGAAG	29259187	29262463
297	R	CTTGAAAAGTGAAGAGATATCCTACAGAAT	29262464	29264721
298	F	GGAGCAGCCAGTAGCCCCAGAAG	29264722	29268159
299	R	CTTGCTTTGCCATTGGACCCTTGTCAG	29268160	29268267
300	F	TGATGCTATCTCTCTTCAAAGGAGGAGAAG	29268268	29270862
301	R	CTTTAAAGCAACAATATGATATGCATCATC	29270863	29272035
303	F	TGGTAAGAAAATGAAATGTAAGGTACGAAG	29277402	29281527
304	R	CTTCCATCTTTTTTCATTTAGAAGCACCAG	29281528	29287941
309	F	GATGGATCTGGAAGGAGGGACACCAAAAG	29292944	29296257
313	R	CTTTAAATTCCTGCTTGTAAAAATTGTTTT	29305044	29309647
314	F	CCAGCCTGTGATGACAGGTGGTAGAAG	29309648	29310307

316	F	ACCTTTCCTCTTCCTAACTGCCATGCAAAG	29310490	29311401
317	R	CTTGACCTCAGTATTCTGTACTCACTCCCC	29311402	29315522
319	F	GCCCCTGAAGATATAGCACAGTCTTGCAAG	29317749	29319160
320	R	CTTGCACGTCTCCATTCATCTTGACTAAAG	29319161	29319272
321	F	TCAAACCATACAGCTGAATATGGAGAGAAG	29319273	29320195
324	R	CTTAAATCTAATAAGATGAAGGAAAATAAC	29321894	29331157
325	F	CACAGCTCCTGCACATCTGGGGAAG	29331158	29332122
326	R	CTTCCAACACCAAATGTGGTTCAATTACGC	29332123	29347103
327	F	CCATTTAAGGCCTGCCCAGCTCAAG	29347104	29352141
330	F	AGATCTGTGCAACTTTCTTCAAGCCACAAG	29356572	29358986
331	R	CTTGGATGCCTTTGATCCCCTTATAACTTA	29358987	29363174
342	F	TCACCACATGTTGAACTCTGGATGGGAAAG	29384407	29389154
345	R	CTTCTACGTAGTAGATTTTTATATGGAATT	29401642	29402954
346	F	TTTGGTTTGGTTTTCTGGGTTGAGGGGAAG	29402955	29404268
349	R	CTTCCCTCCTGGGAATGCTGTGTGG	29406181	29407971
350	F	AGCCTCCAGCTAAATGCCAACAAATGAAAG	29407972	29411033
351	R	CTTAGGATAGGAAGGAATTAGTGATACAGT	29411034	29411488
353	F	CAAATACAAAAAATCTGCCAAACAACAAAG	29411528	29412802
354	R	CTTGGAAACTGCACCATGGAACTCTCACTT	29412803	29413667
355	F	CCCTCTTCTCTTTCAAAGATGGAAAGTAAG	29413668	29419014
359	R	CTTTCTTGAGAAGGAGTTTTCATTCTAGTA	29425367	29425485
361	F	TTACTTATTTGTGACATGGTATCTAAAAAG	29426352	29436153
362	R	CTTCAGGGGAGTGAAAGAATTAAGATTTCT	29436154	29438524
366	R	CTTGCCCCCACAGGGCAGGC	29454397	29457357
367	F	GACCACATCCTGTCTAAACCCTGCCAAG	29457358	29463943
370	R	CTTTGGTATGAGGTAATAAAAATAATTGAA	29466664	29471070
371	F	CTGGCTGTCCTTTGCCACCAAACAAAG	29471071	29483618
372	R	CTTCGTGGTTGTATGCCTGTAATTGTGTTT	29483619	29486276
373	F	GTTCTTTCAGTGATAGATGAGAAGTGTAAG	29486277	29487932
375	R	CTTTCAAGGTTGCAGGTTACAAAATTATTA	29492359	29495456
376	F	TCAAGGAAGGCTACAAGAGGAGAGGCAAG	29495457	29497960
378	R	CTTTGTGAGGACGTGAGTCCGGCTG	29508039	29514012
379	F	GGATATTGTCTCCAATGGTTATGTTAAAAG	29514013	29518402
382	R	CTTTTGGATGCCATTAGCTAGACCTGAGTT	29519437	29521324
383	F	GTTGGCCTGCTATTAATCTGCTTCTACAAG	29521325	29521533
385	R	CTTAACTTACTTCACTAGATTTACTTCTGG	29527218	29527530
386	F	GTAGGCTTGAGCACTTACCAGCAGTGAAAG	29527531	29527778
387	R	CTTTCTACAAACTGGGAAAACCAGCCTTTG	29527779	29530051
388	F	CAGCTGAGAAGACCCAGCACAAATCCAAG	29530052	29530237
390	R	CTTATGAAGTTTTTCAGTTAAAAGTCACAT	29531243	29534953
391	F	AGAGCCACAGTGAACAATGTCTCTAGCAAG	29534954	29536943
393	R	CTTTTACTGGTTTAATATCCTTCCAGGCTT	29537971	29539101
394	F	CTAAAGGTCTTCTTGAGAGGAGCACTGAAG	29539102	29543893
395	R	CTTTACTTTCTGTCACCTTAAAGGTGAAAA	29543894	29545876
399	F	CAGGCTTCTGCCTGCAGAACCAAG	29555730	29559300
400	R	CTTCTCCTAGCACCAACCTTATGATCCTGG	29559301	29560638

402	F	GGAGTAAATGCTACATGACTGCCTGGGAAG	29560665	29560851
403	R	CTTGGTCTTTAATGTAGGCATGATGGGGTA	29560852	29561928
404	F	TAGTTATCCCGTGAGAAAGGTACACTCAAG	29561929	29563972
405	R	CTTGGCAGTTTGATTAATGCATGCTTAAGC	29563973	29566252
407	R	CTTGGGTACCAGTTGCATAAGCGTGG	29567141	29570404
408	F	GACCTGACACTGTTGATGACATTAACAAAG	29570405	29571012
409	R	CTTCCCAACCAAGGTGGGTGGG	29571013	29574785
410	F	CCTCACAACTTCTTTGTAGCAAGTGATAAG	29574786	29577091
411	R	CTTGCGAACAGAATAAAGGACGCATTTACC	29577092	29577561
412	F	ATTATTTTAATGTAGTTAGTAATTTGAAAG	29577562	29579620
414	R	CTTAGCTTCCATTTGTTGGGAAGAGTGGTG	29582421	29582838
415	F	TATATATACTTATGAACATGTTTGTAAAAG	29582839	29584474
417	R	CTTACTTCAGAATTAGGAAAACACAAAGCA	29584537	29585000
418	F	AGCAAAAGCATTAAGTGTAGCAGTGGAAAG	29585001	29590052
419	R	CTTTCCATTCTGAGTCTAGTGACTTAAAGG	29590053	29595530
421	F	TACCCTTATCAGAATGAAGTGTGCACAAAG	29595576	29595777
425	R	CTTGTCACTGCCCATTACCTGACTGTGC	29606038	29607875
426	F	CTATGGATGTTAACCAATAAGCTACAGAAG	29607876	29613408
429	R	CTTCTTATTGGAAAAATTGAAATTTTTCCT	29622314	29623626
431	F	ACATCCATAAGATTATTTCCGATGATCAAG	29626315	29628597
432	R	CTTCATATCTAACATTTGACTCATTGAGAG	29628598	29629851
435	F	TGCTCTACCATGCGTGAAGGTGATATAAAG	29632212	29635239
436	R	CTTAGAACATTAGGAATCATCACTAACTCT	29635240	29637005
437	F	ATAAATCTAACAAACTAAAAATCAGTCAAG	29637006	29638519
438	R	CTTGCTGCCCTCACCTGTGTTGGTAA	29638520	29640675
439	F	CTCATTTCATTTACCTCACTCCTTTAGAAG	29640676	29641133
440	R	CTTGCTTTTGTTGTAGGGATTTTACTTCCT	29641134	29642803
441	F	CATACAAAAGCCAGTTTTCTCAATGTGAAG	29642804	29649071
444	R	CTTGCCACAGCCTAGTTTGAGCCTTAGG	29652798	29652907
447	F	TTCACATGTAACAAAGCCAAATTTATGAAG	29656939	29658283
453	R	CTTGAGTGAGGCTTTATTTTTAGCATTTGG	29677935	29691941
457	F	GAGCCACTGCAGGGCTGGAAG	29701562	29703903
458	R	CTTCAATAAACAAGTGAGGGCGGGTGCA	29703904	29705290
459	F	ACATTTCTGTTGCAGGGCTATTCATGGAAG	29705291	29706210
460	R	CTTAGAATACTTAGTACTTGGTTTACACCT	29706211	29709705
461	F	GGTAACAAGAATAGCATTGAAAATTCTAAG	29709706	29713030
463	R	CTTGGCACTCTTTCAGGAAAGAAAGGACAG	29713093	29716181
464	F	CACCAGCCACCATCAACTACACAAAGAAG	29716182	29720154
465	R	CTTTGGTGTGTAGTACTTTGGAGTGATAAG	29720155	29727114
466	F	CTGAATTTAGGTGTATGTTTACTCAGTAAG	29727115	29730837
468	R	CTTTGGGGATCCTCTGTAAGTGGTTGCTC	29741658	29748329
470	F	CATAATAAAACTGAAAGGAGATGACCCAAG	29750154	29752052
472	R	CTTTCTCAATTGATTCTGAAAGGGAAAATG	29755324	29755456
474	F	TAGTACTTTAGCTAATGGCATTCACTGAAG	29757662	29764922
476	R	CTTCCTCGTTCCAGATGGGGTCTGG	29766525	29768740
477	F	TGAAATGTCCCACAAAGGATTATCAGGAAG	29768741	29769326

478	R	CTTGAGGTTTTCATATCAACAAGGCTCAGT	29769327	29770690
479	F	GGTGAGCTAGTCAGAGCAGTGCTGAAAAG	29770691	29772235
481	R	CTTGTCTTTGAGAGCCCGGTGCCT	29774284	29777533
482	F	TGGCCATTTTATATCTACTTAGGAGAAAAG	29777534	29781541
483	R	CTTTTTCCAAGTCAGGTTAGTAAAAGCAGA	29781542	29782720
484	F	GACCCCAAGGCACCCAACTTCAAG	29782721	29784642
487	F	TTCTGAGCTCCTGTTTTCCCTCAGAGTAAG	29786563	29787575
488	R	CTTATACCTTTATGACCCATAATGCACAGA	29787576	29797222
489	F	GGTTGGAGAGTTTGGAGGCTGAACACAAG	29797223	29805362
490	R	CTTGATTTCTCAAACTACTTATTGATTCGT	29805363	29806188
491	F	GTTTGTTGTGTAGGGGATGTGATCCAAAAG	29806189	29807449
492	R	CTTTCTAAAAAGGGAGGAAGGAAATCCAAG	29807450	29809879
493	F	ATTTGCCACTCAAAATCTGCACTTTCCAAG	29809880	29815593
494	R	CTTCCCTGCTGTAGGGGAGAGCG	29815594	29816516
495	F	GGAGGGGGGAGGTGGGAACAAG	29816517	29820938
496	R	CTTGATGAAGATCTTGGCATGGCAATGCAC	29820939	29835163
497	F	AAGACTCAGTTATACCAATGGTTCAATAAG	29835164	29835981
498	R	CTTCTTTGCGGAATTCCTAGGACGCTAATG	29835982	29836156
499	F	GGTGCAGGGAAAGTTGATAAGGGCAAAG	29836157	29838218
500	R	CTTGCCGCCAGGAAGCTAATTCCTC	29838219	29839237
504	R	CTTTCAGAAATGGAAGACTTTTTTAATTCT	29844732	29845798
505	F	TTCAGCTTTGTCTTGGTGTGTGGACTAAAG	29845799	29846398
506	R	CTTGAATGAGGAAACATAGGCTGAGAGGCC	29846399	29848887
507	F	GTACTGAGCCTAGCAGAGGAAGCTCAAG	29848888	29849213
508	R	CTTTCTACAGCCATGCCATTTACAGAAGCC	29849214	29850560
510	R	CTTCTACCTGCTAACAAACTTTCTCCTGTC	29851352	29854911
511	F	AGTGACCTAGATCTTCCCCTACACTTTAAG	29854912	29855302
512	R	CTTGCGCCTTTCCAGCAGTCTTGAAACATA	29855303	29859544
513	F	CATCATCTGTGCCATTTCTTAGCCCCAAAG	29859545	29863829
514	R	CTTATCCACCATGGCCCCAAGATTATCTTT	29863830	29864649
517	F	CTCTGAAACATGCAACCACACAGGCAAAG	29870040	29874795
518	R	CTTAACCAAGTAACACTACCAACTGGAGAC	29874796	29884457
521	F	TTTTTTTTCCCATTAGGCGAAATTTTAAAG	29896489	29896823
522	R	CTTCATCTGACCTCGTTTGCACGAAGCTC	29896824	29897292
523	F	TAAAACTTGCAGCAACTTTCCTGTGTTAAG	29897293	29900118
525	F	AATATGGTTTCCTTATGATGTAGCTTTAAG	29902876	29906098
526	R	CTTCCAAGTTTCTGAAGCATCCTCACCAGA	29906099	29907124
527	F	TTGTGATCAAATTTACATGTCTAAGCTAAG	29907125	29909799
529	R	CTTTTGACTAGTGGACTTTATCTGCTCTCA	29911938	29913290
530	F	TTCCTTTAGCTTTATATATTGATGGAGAAG	29913291	29913607
532	R	CTTCAGTTATACAAAGTTCTTTCATGCCCC	29913932	29918392
533	F	AGATCTCTCTACTTGTTGATTAATAGCAAG	29918393	29927218
535	R	CTTAGGTATAAAAGAAAATCTTTAAATACC	29931042	29938025
536	F	AAGTTAATTGAATGAAAATGATCAACTAAG	29938026	29947212
537	R	CTTTCTGACCAAAGTTGACAACAGCACTAT	29947213	29949609
538	F	CATGTGTTGTAGAGGCTGGGAACCACAAG	29949610	29953599

539	R	CTTTTGTTTATGGTCTGGAATGTACCCATT	29953600	29958998
540	F	CATGTGGGTTAACAGTGAGTTAAGCCCAAG	29958999	29960405
542	R	CTTCAGCCCCTTGCCTTGATGCC	29960412	29963242
543	F	TAGACATGTGAATTTTATTCTTGGAGGAAG	29963243	29963692
544	R	CTTTGAAAGACCCATGTCATAGTACGTGTT	29963693	29963925
546	F	GATTTTTTTTTCCTAGGGGTTTATTTCAAG	29967853	29970155
547	R	CTTAGGCTCAAAGATGACTGCAGAGGAGAG	29970156	29977864
548	F	AAGGAAAAAAAAAAGAATAAGGTTTCAGAAG	29977865	29978715
549	R	CTTGCAAGCTAGCTAGCCCAAGGGATAC	29978716	29979319
550	F	TGTACAGTATAAGCACCCCTATAGCCAAAG	29979320	29988808
551	R	CTTTCAATTTTTCCACATCCTGCCCAACAC	29988809	29991479
553	R	CTTTCCAGTATTGCTGGAGCTGTAGTCCGA	29995277	29996459
554	F	GCAACTCAGTGAAATTAACCAAAGATGAAG	29996460	30001398
555	R	CTTTCCAGGCATAAAGCAGAGACAGGCAAA	30001399	30005000

***USP22* (chr.11; mm9)**				
1	R	CTTTGAGACTAGACCGAAGTCTCCAGAATC	60917307	60918190
2	F	GGGAGGAAGATAAAAAGATGGGGATGGAAG	60918191	60924238
5	F	CTGTCAGTCTCCTGCTGCCACTAACAAG	60932406	60937638
6	R	CTTGAGGAGACAGCACTGCTGGTAGATAG	60937639	60942991
9	F	GTGTCAAGGAGGCAGACCTTCAGGAAG	60949684	60966226
10	R	CTTACTGTGGGCTACCCATTTGTACTCTTA	60966227	60970541
11	F	AACCCATCTCTAACAATCCCTTTGTAAAAG	60970542	60976197
12	R	CTTTTATTTATAAGAGATCTTAGCTAATGA	60976198	60977387
14	F	GCAGGAGTCTAAGCCACCAGGGAAG	60977797	60977935
16	R	CTTTTTTCCAAATAGTAACCCACAGGGCCA	60981479	61000571
17	F	ACTTTTCCCAAAATCCAAGCTGACTTCAAG	61000572	61003268

Table S7. 5C sequencing reads

E11.5 tissues

Sample	Number of reads	Number of used reads
Limb	1438064	677006
Limb-replicate	2483452	1715790
Body	1661514	1057871
Head	1815022	1050062
Head-technical replicate	1771610	932274

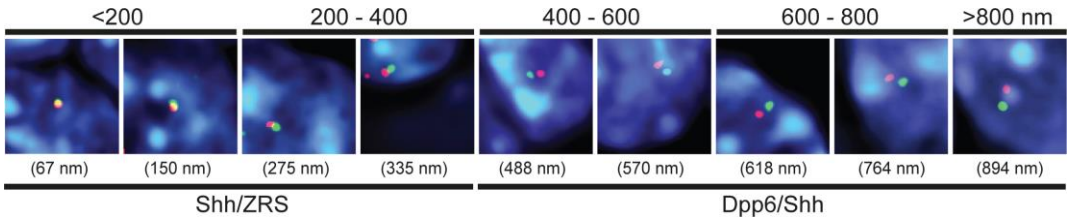

Figure S1. Images of representative nuclei showing probe pairs at various distances apart. *Shh*/ZRS probes up to 400 nm, *Shh*/*Dpp6* probes shown for distances greater than 400 nm. All nuclei are from E10.5 ZPA.

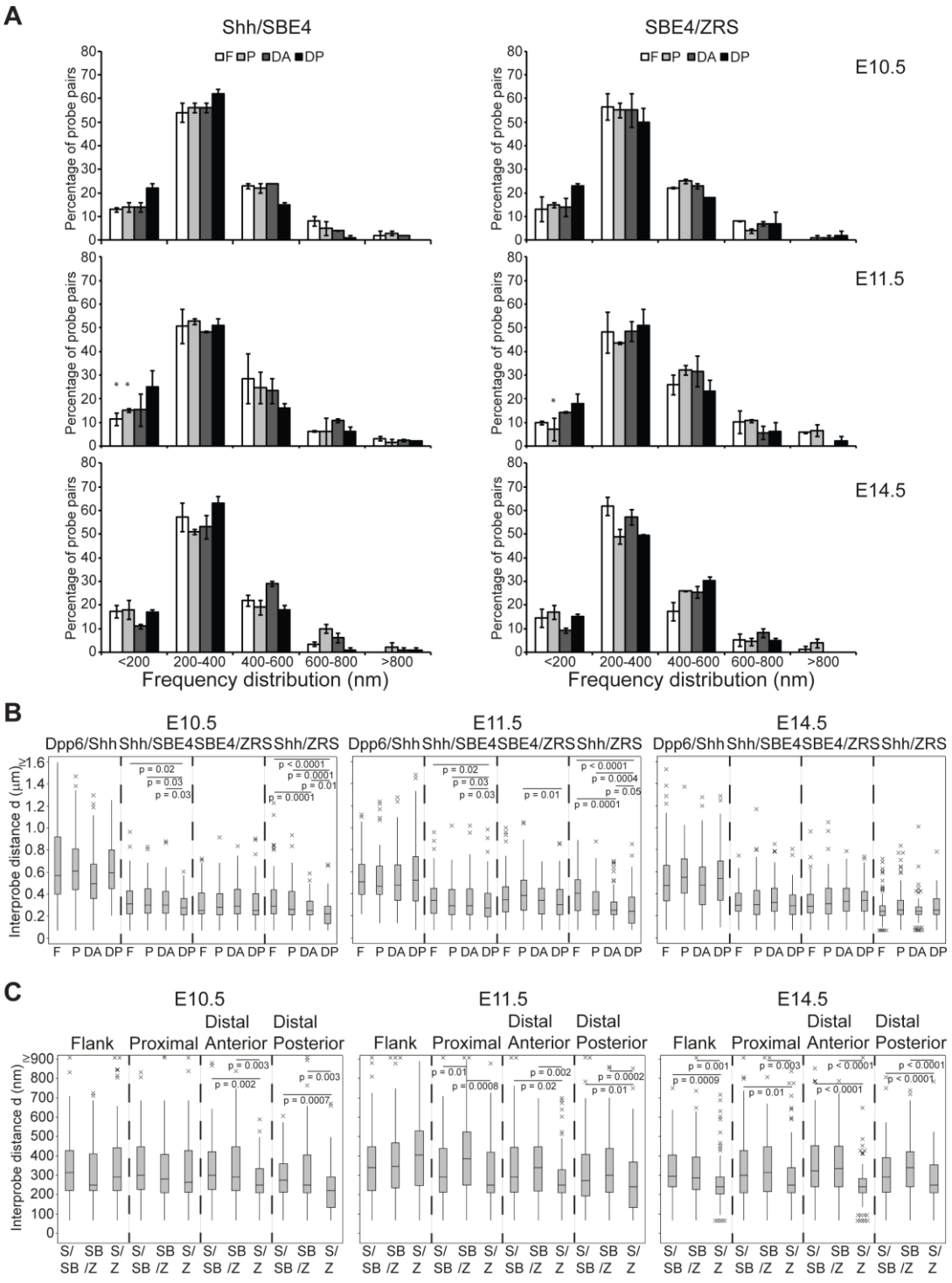

Figure S2. Spatiotemporal frequency distribution of Shh/SBE4 and SBE4/ZRS interprobe distances, and spatiotemporal interprobe distances of *Shh*/*Dpp6*, *Shh*/SBE4, SBE4/ZRS and *Shh*/ZRS. (**A**) Frequency distributions of interprobe distances (d) in 200 nm bins between *Shh* and SBE4 probes, and SBE4 and *ZRS* probes, in proximal and distal regions of the murine forelimb bud and adjacent flank at E10.5, E11.5 and E14.5 temporal stages. F: flank, P: proximal limb, DA: distal anterior limb, DP: distal posterior limb (ZPA in E10.5 and E11.5 sections). n = 70 – 100. Error bars represent SEM obtained from two or three different tissue sections. The statistical significance between data sets was examined by Fisher's Exact Tests: * $p < 0.05$. (**B**) Boxplots show the distribution of interprobe distances (d) in micrometres between the four sets of probe pairs in E10.5, E11.5 and E14.5 distal anterior and posterior and proximal forelimb, and flank tissue. The statistical significance between data sets was examined by Mann-Whitney U Tests. (**C**) Boxplots comparing the distribution of interprobe distances (d) in nanometres of the three sets of probe pairs located across the *Shh* regulatory region in E10.5, E11.5 and E14.5 distal anterior and posterior and proximal forelimb, and flank tissue. S/SB: *Shh*/SBE4, SB/Z: SBE4/ZRS, S/Z: *Shh*/ZRS. The statistical significance between data sets was examined by Mann-Whitney U Tests.

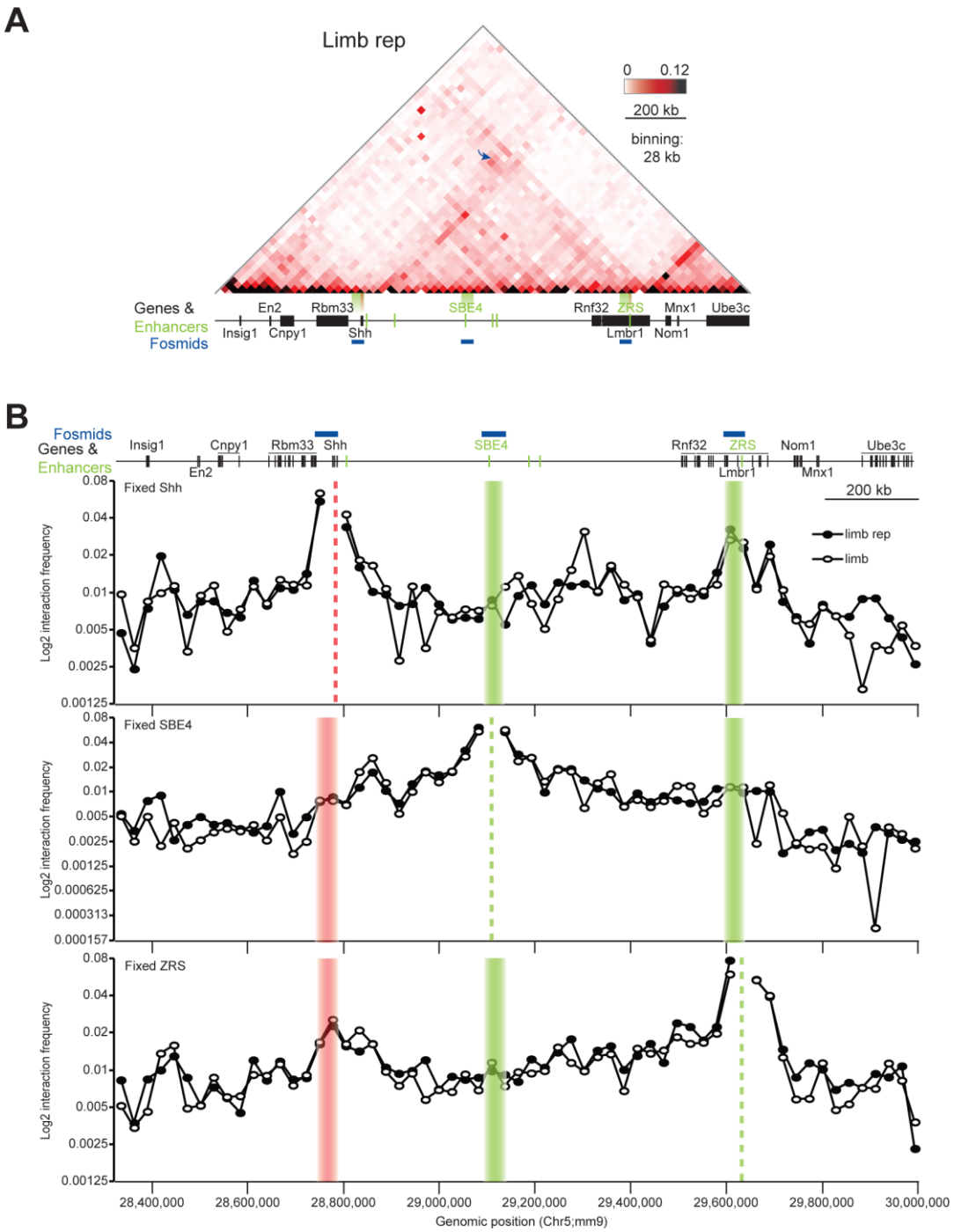

Figure S3. 5C-seq enriched interactions between *Shh* and ZRS are recapitulated in a biological replicate. (**A**) 5C heatmap shows the average interaction frequencies (28-kb bins) across *Shh* and its regulatory domain in E11.5 limb bud cells (replicate biological sample). Arrows indicate interaction frequencies between windows containing Shh and ZRS. Interaction frequencies are colour-coded according to the corresponding scales as described in Figure 4A. (**B**) Virtual 4C analysis obtained by extracting 5C interactions with viewpoints fixed at *Shh*, SBE4, and ZRS. Dashed lines indicate the position of the fixed viewpoint from the *Shh* genomic region (orange) or regulatory/structural elements (green). Data from limb bud and limb bud replicate cells are in open and filled circles, respectively.

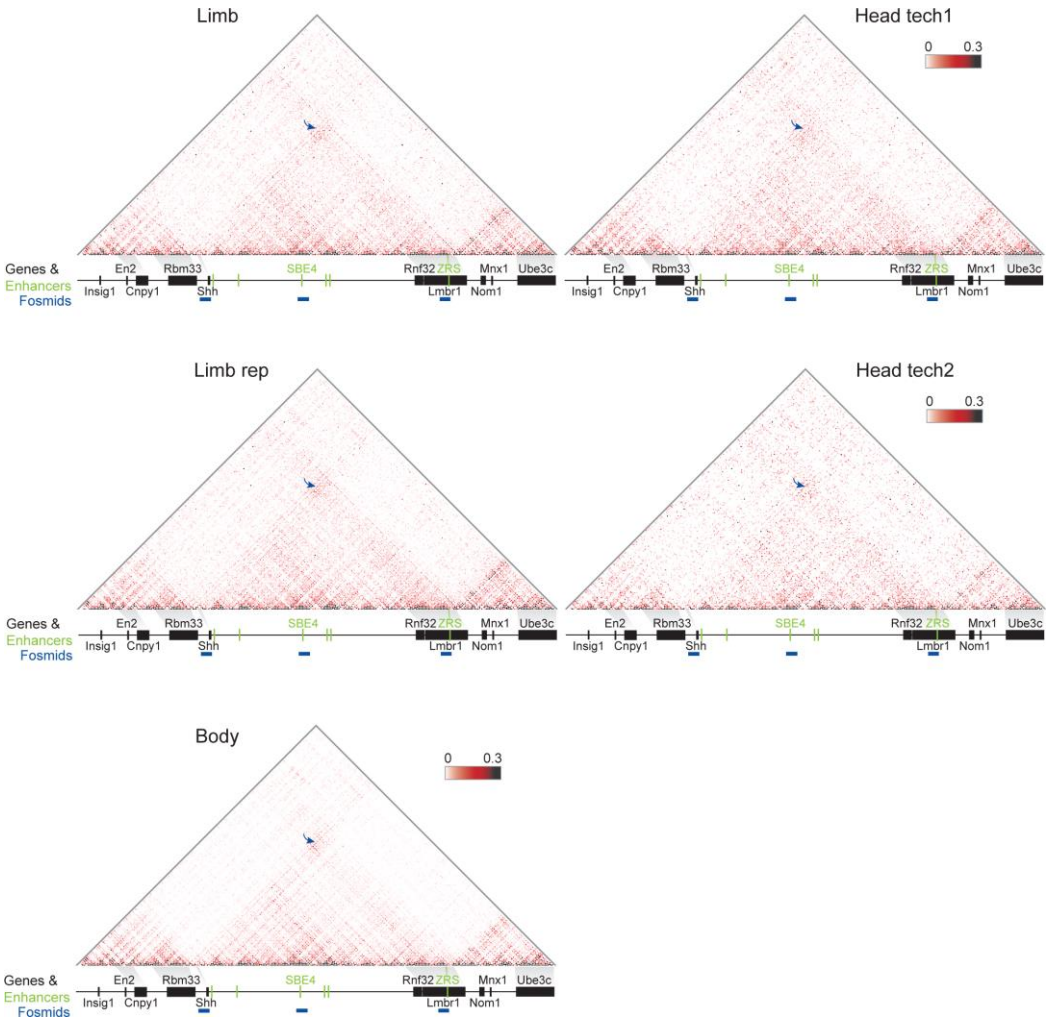

Figure S4. Unprocessed heatmaps of the normalised E11.5 limb bud, body and head cells 5C datasets. 5C-seq normalised data are presented in the heatmap form according to colour scales as described in Figure 4A. Genes are indicated in black, regulatory elements in green and fosmid probes in blue. Grey shading highlight the position of the genes in the 5C heatmaps. Blue arrows indicate enriched interactions between genomic fragments over the *Shh* locus and ZRS locus.